The University of Chicago Science Series

ADVANCES AND APPLICATIONS OF MATHEMATICAL BIOLOGY

NICOLAS RASHEVSKY

THE UNIVERSITY OF CHICAGO SCIENCE SERIES

*

THE UNIVERSITY OF CHICAGO SCIENCE SERIES, established by the Trustees of the University, owes its origin to a belief that there should be a medium of publication occupying a position between the technical journals with their short articles and the elaborate treatises which attempt to cover several or all aspects of a wide field. The volumes of the series will differ from the discussions generally appearing in technical journals in that they will present the complete results of an experiment or series of investigations which previously have appeared only in scattered articles, if published at all. On the other hand, they will differ from detailed treatises by confining themselves to specific problems of current interest, and in presenting the subject in as summary a manner and with as little technical detail as is consistent with sound method.

Advances and Applications of Mathematical Biology

THE UNIVERSITY OF CHICAGO PRESS
CHICAGO, ILLINOIS

*

THE BAKER & TAYLOR COMPANY
NEW YORK

THE CAMBRIDGE UNIVERSITY PRESS
LONDON

THE MARUZEN-KABUSHIKI-KAISHA
TOKYO, OSAKA, KYOTO, FUKUOKA, SENDAI

THE COMMERCIAL PRESS, LIMITED
SHANGHAI

Advances and Applications of Mathematical Biology

By
NICOLAS RASHEVSKY
The University of Chicago

The University of Chicago Press
Chicago, Illinois

COMPOSED AND PRINTED BY THE UNIVERSITY OF CHICAGO PRESS
CHICAGO, ILLINOIS, U.S.A.

TO MY WIFE

PREFACE

In our previous book, *Mathematical Biophysics: Physicomathematical Foundations of Biology*, we attempted to present a system of mathematical biology based on physical concepts. While we tried, wherever it was possible, to establish contact between mathematical conclusions and experimental facts, still the theoretical foundations occupied a much more prominent place than their actual applications.

During the two years which have elapsed since the completion of the manuscript of that book, considerable progress has been made in this field, especially in the direction of applications of the mathematical theory to various observations. This progress has largely been due to the use of the new approximation method, very briefly outlined in the Appendix of the former book. In the present book it is used as the foundation for all the developments of chapters i–vi. Many problems, which offered very little hope for a solution in the near future by means of the standard methods used before, have found their solutions by means of the new approximation method. The latter also considerably simplifies the mathematics used and will, we hope, make the present book accessible to a much wider circle of readers.

Even more than the previous book, the present one owes a very large amount of most important material to the work of others. Without the important researches of A. S. Householder, H. D. Landahl, Robert R. Williamson, Alvin M. Weinberg, and Gale Young this book would never have been possible. Throughout the volume the reader will find numerous references to the work of the above-mentioned members of the Chicago group of mathematical biophysicists.

Yet, in spite of the progress made, the author is perfectly

well aware that he is not presenting anything like a finished product. This book is to be regarded merely as an early milestone on the road to further progress.

The year 1939 has witnessed another important progress in the field of mathematical biology, namely, the founding of a journal devoted exclusively to publications in that field. *The Bulletin of Mathematical Biophysics* is published quarterly by the Psychometric Corporation under the editorship of the author of this volume and is entering now upon its second year of existence. The editorial and publication offices are located at the University of Chicago, 5822 Drexel Avenue, Chicago, Illinois. Practically all the original publications in mathematical biophysics appear in the *Bulletin*, which thus constitutes a record of research in that field.

The author's wife, Mrs. Emily Rashevsky, to whom this book is dedicated, has been responsible for the preparation of the whole manuscript. For the painstaking and tedious work of checking all formulas, both in the manuscript and in proofs, the author is indebted to Dr. Alston S. Householder, Mr. Herbert D. Landahl, Dr. Alvin M. Weinberg, and Mr. Gale Young. Thanks are due Mr. Robert R. Williamson for the preparation of the Index, and Mr. Robert Dubin for the preparation of a number of original drawings. Furthermore, the author is under obligation to Professor N. J. Berrill for permission to reproduce Figures 2 and 3 from *Growth*, 1, 213, 1937; to Professors A. V. Hill and D. Y. Solandt for permission to reproduce Figures 29 and 30 from *Proceedings of the Royal Society of London*, B, 120, 400, 1936, and 121, 113, 1936; to Professor C. J. van der Klaauw for permission to reproduce Figures 12, 13, 14, and 26 from *Acta biotheoretica*, A, 4, Part II, 1938; and to the Psychometric Corporation for permission to reproduce Figures 39, 41–46, 48, 49, and 53–58 from different papers published in *Psychometrika*. He also wishes to express his thanks to Professor Robert Chambers for permission to

use for measurements a part of his micromotion film of dividing blastomeres.

Finally, acknowledgment must be made to the University of Chicago Press and particularly to Miss Mary D. Alexander, of the Editorial Department, for unfailing co-operation during the process of manufacturing the book. This co-operation must have been rendered especially difficult by the author's tendency to make minor improvements and last-minute alterations at the time of proofreading.

N. RASHEVSKY

CHICAGO, ILLINOIS
January 2, 1940

TABLE OF CONTENTS

CHAPTER I

DIFFUSION PHENOMENA IN CELLS

In biology, as in other natural sciences, we find that, in spite of the tremendous variety of phenomena observed, certain uniformities become apparent upon closer examination. The existence of such uniformities, of such a *unity in diversity*, is one of the prerequisites for the possibility of existence of any science. Among other uniformities found in biology, one of the most conspicuous is the mode of propagation of life on earth. While the methods of multiplication of organisms are almost infinite in their variety, they all reduce, in the ultimate analysis, to cell division. Again, the details of cell division vary greatly from case to case. But everywhere the *fundamental* aspects of the phenomenon is the same: a cell, at a certain stage of its life, spontaneously breaks up, divides into two or more cells, which grow further as individual living units. It is in keeping with the spirit of all exact sciences to look for a simple common cause of such a general phenomenon. To effect such a division, some force of unknown origin must exist. The very commonness of the cell division suggests that there must be some common cause underlying all cases of cell division and responsible for the force that produces it, though additional factors may strongly modify the manifestations of cell division from case to case. However, if we assume such a common cause, it is only natural to connect it with some other phenomenon, one at least just as common to all living cells. Moreover, since the phenomena of spontaneous division are almost exclusively the prerogative of the living, we must therefore, to find their cause, look for something that is especially characteristic of the living. Electrical phenomena,

which have sometimes been advocated for the explanation of cell division, do not meet that last criterion. They are universal, found equally frequently in the living and the non-living. This, however, does not imply that their *secondary* effects upon cell division may not be of prime importance.

One phenomenon, common to all growing and dividing cells, suggests itself particularly. In every cell there is always present a flow of different substances. Metabolism is one of the things common to all living organisms—and almost exclusively common to them. Various foodstuffs flow into the cell; different waste products flow outward. The principal agency of this transport of substances is diffusion. When, for instance, a substance is produced inside a cell, its concentration there exceeds the concentration in the surrounding medium. Because of the thermal agitation of its molecules, the substance is transported from regions of higher concentrations to those of lower concentrations. This transport would be very rapid if it were not for the resistance offered by the medium, in which the substance is dissolved, to the movement of the molecules flowing through it. But if the medium exerts a force of resistance opposing the flow, then the flowing molecules, by virtue of Newton's Third Law, exert a force on the medium, in the direction of flow. Hence, in every living cell, by virtue of its ever present metabolism, there always is a more or less complex system of forces acting upon different parts of the cell. And it behooves us to investigate whether the magnitude and direction of these forces may not be sufficient to produce eventually a division of a cell.

Since these forces are due to diffusion flow, we must, as a preliminary study, investigate various general properties of such a flow. The direct mathematical manner of studying such a problem is to investigate the solutions of the differential equations of diffusion, well known in physics. Here, however, we meet at once with a serious difficulty. Given the size

and shape of a cell, the rates of production or consumption of a given substance, the diffusion coefficients, which characterize the resistance of the medium to the flow of these substances, and some other physical constants, we can determine the resulting flows *in principle*. But actually, except for the very simplest cases, the mathematical difficulties attending the solutions of the differential equations are, at present, insuperable.

Moreover, even granting that we could solve exactly the differential equations of diffusion for a large number, or even for all cases, this would still leave us with a rather serious handicap. The distribution of the diffusion flows in every individual case depends, among other things, on the exact shape of the cell. A slight variation of the latter will modify the analytical expressions describing the distribution of concentrations and flows. But, since there are no two cells perfectly alike, the exact solution of the problem for a given case would contain a tremendous amount of detail which is biologically insignificant because it applies *only to the given case*. If we wish to draw general conclusions, we shall have to drop most of the details and pick out only the essential features of the solution, which would hold for a large number of different but somewhat similar cases. But with a complicated mathematical solution this cannot be done unless we actually make a study of a very large number of cases and compare the results. This certainly would not be availing ourselves of the powers of generalization offered by mathematical analysis.

Two ways out of this difficulty may be suggested. One is to solve the problem *exactly* for some simple case and to trust that the conclusions so reached will hold, with some degree of approximation, for those actual and more complex cases which do not deviate too much from the ideal simple case chosen. This way has been followed in mathematical biology hitherto. Inasmuch as the sphere is almost the only shape for

which the general differential equations of diffusion can be solved exactly, we have limited ourselves to the theoretical study of spherical cells. In spite of such a limitation, some rather interesting and important results have been obtained (see our *Mathematical Biophysics* [Chicago: University of Chicago Press, 1938]; hereafter referred to as "*MB*"). While the greatest majority of cells do not possess a spherical shape, still a rather large class of cells are almost spherical, and we would naturally expect that for such cells our mathematical deductions will hold approximately. But the inadequacy, in the long run, of such a procedure is obvious, especially where we are dealing with the cell division, during which the shape of a cell is anything but spherical. We could, by comparing the initial stage of a large spherical cell with the end-stage of two spherical half-sized cells, demonstrate that the energy relations are such that the end-stage is mechanically more stable than the initial stage. From that we would argue that at a given size the large cell may spontaneously divide in two. But nothing could be said about the most important intermediate stages and about the mechanism proper of division.

The second method of approach, which has been outlined in the Appendix of our previous book (*MB*), attacks the difficulty in a radically different way. We do not attempt to solve the mathematical problems exactly for a shape specified in detail, but rather we confine ourselves to finding mathematical relations between such gross features as are common to all cells of a given type, in spite of a difference in detail. We may have two cells, with a given approximate ratio of length to width. The biologist would describe them as belonging morphologically to the same class. Yet the detailed shape of their surfaces will differ enormously from a mathematical point of view. It would be quite illusory, and practically unnecessary, to take into consideration such small, biologically irrelevant details. But, inasmuch as two such cells behave biologically

in a similar way, we must look for such relations as depend only on their gross shape and structure. It seems to be more promising, from a practical point of view, to deal with relations involving only orders of magnitudes, and not with exact expressions.

Let us consider a cell having an over-all approximate length $2r_1$ and an over-all width $2r_2$ (Fig. 1). Its internal structure may be inhomogeneous; but for the time being we shall neglect those inhomogeneities, for in detail they will differ from cell to cell. Consider first, for simplicity, that a substance is

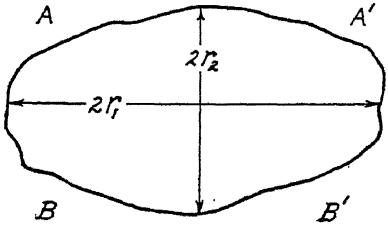


FIG. 1

produced inside such a cell, at an *average* rate q gm cm^{-3} sec^{-1}, a rate which does not depend on the concentration of the substance. The actual rate of production will undoubtedly vary from point to point within the cell. But the details of such variations again constitute *individual* characteristics. Whatever we measure and observe are only average values, and only for them may we find some general regularities.

Similarly, let us consider the average concentration $\bar{c}$ of the substance inside the cell. We shall express concentration in gm cm^{-3}. Again neglecting possible detailed individual variations, all we can say is that, since the substance diffuses outward and flows from center to periphery, therefore the actual concentration will be higher around the center and will drop toward the periphery. The average value $\bar{c}$ will be found approximately somewhere *midway between periphery and cen-*

ter. Along the periphery the concentration will also not be constant. Neglecting detailed variations, let us denote the average concentration inside the cell at the periphery along the "ends" AB and $A'B'$ (Fig. 1) by c_1, and the average peripheral concentration along the "sides" AA' and BB' by c_2. Then the average drop of concentration per unit length is

$$\frac{\bar{c} - c_1}{r_1} \tag{1}$$

in the longitudinal direction and

$$\frac{\bar{c} - c_2}{r_2} \tag{2}$$

in the transversal direction.

If D_i denotes the coefficient of diffusion for the produced substance inside the cell, then, according to the fundamental law of diffusion, the average flows of the substance in the longitudinal and transversal directions are equal, respectively, to

$$2D_i \frac{\bar{c} - c_1}{r_1} \text{ gm cm}^{-2} \text{ sec}^{-1}$$

and (3)

$$2D_i \frac{\bar{c} - c_2}{r_2} \text{ gm cm}^{-2} \text{ sec}^{-1}.$$

When a cell produces a substance which diffuses outward, then in the immediate neighborhood of the cell there will be an excess of that substance in the external medium. If, as will generally be the case, the external medium, which we consider as extending infinitely, contains in the absence of the cell the substance in question in a concentration c_0, then in the vicinity of the cell the concentration of that substance will

be greater than c_0. For an oblong cell this external concentration near the cell will, in general, be different along the "sides" AA' and BB' and along the "ends" AB and $A'B'$ (Fig. 1). Let us denote this concentration at the "ends" by c_1' and at the "sides" by c_2'.

If the cell is bounded by a membrane of some sort, then the flow of the substance through this membrane is approximately proportional to the difference of concentrations on both sides of the membrane, as is known from observations on membrane permeabilities. If h denotes the permeability of the membrane, then the flow per second per square centimeter of the membrane is, for the "ends" and "sides," equal, respectively, to

$$\left.\begin{array}{l} h(c_1 - c_1') \text{ gm cm}^{-2} \text{ sec}^{-1} \\ \text{and} \\ h(c_2 - c_2') \text{ gm cm}^{-2} \text{ sec}^{-1}. \end{array}\right\} \qquad (4)$$

From expressions (3) and (4) it follows that the dimension of the diffusion coefficient is length^2/time, while that of permeability is length/time. In absolute units, therefore, the diffusion coefficient is expressed in cm^2 sec^{-1}, while the permeability is expressed in cm sec^{-1} (*MB*, p. 11).

The excess of the substance outside the cell becomes less and less as we move farther away from the cell. Practically, it is appreciable only over a finite range of distance from the boundary. We shall denote this distance by δ. The distance δ is of the order of magnitude of the over-all size of the cell. The justification of this assertion will become apparent subsequently from comparison of our approximate equations with some known exact ones. The average drop of concentration per unit length outside the cell in its neighborhood is then equal, in the longitudinal direction, to $(c_1' - c_0)/\delta$ and, in the transversal direction, to $(c_2' - c_0)/\delta$. Denoting by D_e the

diffusion coefficient in the external medium, we have, for the corresponding flows,

$$\left.\begin{aligned} &\frac{D_e(c_1' - c_0)}{\delta} \text{ gm cm}^{-2} \text{ sec}^{-1} \\ \text{and} \\ &\frac{D_e(c_2' - c_0)}{\delta} \text{ gm cm}^{-2} \text{ sec}^{-1}. \end{aligned}\right\} \tag{5}$$

The amounts given by expressions (3) arrive per second to each square centimeter of the membrane. They must be equal to the corresponding amounts passing through every square centimeter of the membrane, as given by the expression (4). These, in their turn, must be equal to the corresponding amounts leaving every square centimeter on the outside of the membrane and given by expressions (5). Hence, writing down those required equalities, we obtain the following four equations:

$$2D_i(\bar{c} - c_1) = r_1 h(c_1 - c_1'), \tag{6}$$

$$2D_i(\bar{c} - c_2) = r_2 h(c_2 - c_2'), \tag{7}$$

$$\frac{2D_i}{r_1}(\bar{c} - c_1) = \frac{D_e}{\delta}(c_1' - c_0), \tag{8}$$

$$\frac{2D_i}{r_2}(\bar{c} - c_2) = \frac{D_e}{\delta}(c_2' - c_0). \tag{9}$$

The total area of the "end" surfaces, AB and $A'B'$, is of the order of magnitude of $2\pi r_2^2$. Therefore, the total amount of substance leaving through the ends is, because of expressions (4) and (6), equal to

$$\frac{4\pi D_i r_2^2(\bar{c} - c_1)}{} \text{ gm sec}^{-1} \tag{10}$$

The area of the "sides" AA' and BB' being of the order of magnitude of $4\pi r_1 r_2$, the total amount leaving through the "sides" is equal to

$$8\pi D_i r_1(\bar{c} - c_2) \text{ gm sec}^{-1}. \tag{11}$$

The volume of the cell is approximately $\frac{4}{3}\pi r_1 r_2^2$. If we approximate it by a cylinder, we would get $2\pi r_1 r_2^2$, which is of the same form except for the coefficient. For shapes with relatively rounded-up ends the first expression offers a somewhat better approximation. However, since we are after general types of relations rather than after exact numerical values, which differ from cell to cell anyway, it does not matter which expression we choose. This point shall be further illustrated below, when we discuss some results. For definiteness we choose the first expression for the volume. Then the total amount of substance produced in the cell per second is given by

$$\tfrac{4}{3}\pi r_1 r_2^2 q \text{ gm sec}^{-1}, \tag{12}$$

where q is the rate of production per cubic centimeter. The total amount of substance present inside the cell is equal to $\frac{4}{3}\pi r_1 r_2^2 \bar{c}$ gm, and its rate of change with respect to time is equal to

$$\tfrac{4}{3}\pi r_1 r_2^2 \frac{d\bar{c}}{dt} \text{ gm sec}^{-1} \tag{13}$$

This total rate of change is equal to the total amount produced per second less the total amount leaving the cell per second. Hence, expression (13) must be equal to expression (12) less the sum of expressions (10) and (11). Writing down this requirement, shortening both sides by $\frac{4}{3}\pi r_1 r_2^2$, and factoring out D_i, we find

$$\frac{d\bar{c}}{dt} = q - 3D_i\left(\frac{\bar{c} - c_1}{r_1^2} + 2\,\frac{\bar{c} - c_2}{r_2^2}\right). \tag{14}$$

Eliminating c_1' from equations (6) and (8), we obtain, after very simple rearrangements,

$$c_1 = \frac{2(D_iD_e + \delta D_i h)\bar{c} + r_1 h D_e c_0}{2D_iD_e + 2\delta D_i h + r_1 h D_e}. \tag{15}$$

Similarly, eliminating c_2' from equations (7) and (9), we find

$$c_2 = \frac{2(D_iD_e + \delta D_i h)\bar{c} + r_2 h D_e c_0}{2D_iD_e + 2\delta D_i h + r_2 h D_e}. \tag{16}$$

From (15) and (16) we have

$$\left.\begin{aligned} \bar{c} - c_1 &= \frac{r_1 h D_e}{2D_iD_e + 2\delta D_i h + r_1 h D_e}(\bar{c} - c_0) \\ \text{and} \\ \bar{c} - c_2 &= \frac{r_2 h D_e}{2D_iD_e + 2\delta D_i h + r_2 h D_e}(\bar{c} - c_0). \end{aligned}\right\} \tag{17}$$

Introducing expressions (17) into equation (14) and using the abbreviation

$$\left.\begin{aligned} \Lambda = &\frac{r_1 r_2}{3hD_iD_e} \times \\ &\frac{(2D_iD_e + 2\delta D_i h + r_1 h D_e)(2D_iD_e + 2\delta D_i h + r_2 h D_e)}{2(2D_iD_e + 2\delta D_i h + r_1 h D_e)r_1 + (2D_iD_e + 2\delta D_i h + r_2 h D_e)r_2}, \end{aligned}\right\} \tag{18}$$

we find that equation (14) may be written thus:

$$\frac{d\bar{c}}{dt} = q - \frac{\bar{c} - c_0}{\Lambda}. \tag{19}$$

Equation (19), integrated in the usual way, gives

$$\bar{c} = c_0 + \Lambda q - C\Lambda e^{-t/\Lambda}, \tag{20}$$

where C is a constant of integration, determined by the initial conditions, that is, by the value which $\bar{c}$ has at a given fixed moment. Quite regardless of the value of C, and therefore regardless of the initial conditions, the last term of equation (20) decreases to zero with increasing t. Hence, after a sufficient lapse of time, $\bar{c}$ is always given by

$$\bar{c} = c_0 + \Lambda q \,, \tag{21}$$

in which Λ is given by equation (18) and is independent of time.

When we make $r_1 = r_2 = r_0$, which corresponds to a rounded-up shape of the cell, expression (18) for Λ simplifies considerably. Making this simplification and introducing the result into equation (21), we find

$$= c_0 + \left(\frac{2}{9}\frac{r_0}{h} + \frac{r_0^2}{9D_i} + \frac{2}{9}\frac{r_0\delta}{D_e}\right. \tag{22}$$

We may compare this expression with the one obtained by solving exactly the differential equation of diffusion for a perfect sphere. In this case the average value $\bar{c}$ is obtained by integrating the right-hand side of equation (22) of chapter ii of *MB* over the whole volume of the sphere and then dividing it by the volume. We thus find

$$\bar{c} = c_0 + \left(\frac{r_0}{3h} + \frac{r_0^2}{15D_i} + \frac{1}{3}\frac{r_0^2}{D_e}\right) q \,. \tag{23}$$

We see that equations (22) and (23) are the same form if we make $\delta = r_0 \sim r_1 \backsim r_2$, which justifies the assertion about δ made on page 7. The coefficient of the first term in parentheses in equation (23) is somewhat larger than in equation (22), but the other two coefficients are smaller.

Another interesting comparison is made by considering the

general case when $r_1 > r_2$, and when both D_e and h are very large, so that we may put in equation (18) $D_e = h = \infty$. The expression for Λ again simplifies, and equation (21) now gives

$$\bar{c} = c_0 + \frac{1}{3D_i} \frac{r_1^2 r_2^2}{2r_1^2 + r_2^2} q . \tag{24}$$

Gale Young[1]* has computed the *exact* expression for the average concentration inside an ellipsoid of revolution with semi-axes r_1 and r_2 for the case $D_e = h = \infty$. His expression is

$$\bar{c} = c_0 + \frac{1}{5D_i} \frac{r_1^2 r_2^2}{2r_1^2 + r_2^2} q , \tag{25}$$

which again is identical with equation (24) except for a difference in the numerical coefficient.

As we have emphasized above, the individual variations from cell to cell are so large that it would be futile to compare exact numerical values. But the regularities that are found in the behavior of cells, in spite of these individual variations, suggest a search for relations of a general form which may remain invariant from cell to cell despite variations of numerical values. Here we are dealing with precisely that kind of relations. It will also be more clear now as to what we meant on page 9 by saying that it would not matter much whether we choose one or another approximation for the volume of the cell. A different choice in that case would merely modify the numerical value of the coefficients in equations (22) and (24) without changing their form.

It is, of course, justifiable, to some extent, to make an approximate mathematical study of oblong cells by assuming that their properties will be closely enough described by those of ellipsoids for which in some cases the mathematical prob-

* All numbered references are given at the end of each chapter.

lem can be solved exactly. But the difference between an actual cell and an ellipsoid is so great that we shall be able practically to apply only the most general properties of the solution anyway, and the mathematical exactness becomes rather superfluous. If, moreover, we compare the derivation of our equation (24) with that given by Gale Young[1] for equation (25), we shall see at once the labor- and time-saving advantages of the approximation method. Moreover, it may be noted that, while equation (21), together with equation (18), gives us an approximate solution for the most general case, when D_i, D_e, and h are finite, the ellipsoidal problem has, so far, been solved exactly only for $h = \infty$.

We derived here the equation (21) for the stationary state as a particular case of the general equation (19). But we can just as well derive equation (21) directly, without integrating a differential equation, by limiting ourselves, from the outset, to stationary states. In such a case we require that the total amount of substance produced per second in the cell equals the total amount leaving the cell per second. This requirement gives us, instead of equation (14), the following one:

$$q = 3D_i\left(\frac{\bar{c} - c_1}{r_1^2} + 2\,\frac{\bar{c} - c_2}{r_2^2}\right). \tag{26}$$

Introducing into this equation the expressions (17) and using again the abbreviation (18), we at once obtain equation (21). Thus, for a stationary state the whole problem reduces to the solution of five simple linear algebraic equations, namely, (6)–(9) and (26). For the general nonstationary case we have to integrate a very simple ordinary differential equation (19). The exact solution of a diffusion problem requires, in general, the handling of partial differential equations.

If, instead of a substance produced, we consider a substance consumed by the cell and flowing into it, we obtain the same

expressions, except that we must substitute $-q$ for q (*MB*, p. 9). Since Λ is always positive, equation (21) shows that for a produced substance there is always an excess of substance inside of the cell, as compared with the outside, while for a consumed substance there is a deficiency. Since $\bar{c}$ cannot be negative, therefore for $q < 0$, equation (21) has a meaning only for $c_0 > \Lambda |q|$, where $|q|$ is the absolute value of q (cf. *MB*, chap. iii).

Equation (20) shows that the smaller the Λ, the more rapidly the exponential term vanishes and the more rapidly the stationary state is reached. The expression Λ decreases with increasing h, D_i, and D_e and increases with increasing r_1 and r_2. This can be seen directly either from expression (18) or, still easier, from the special cases leading to expressions (23) and (24). Thus, in a sense, the quantity Λ measures the resistance offered to the diffusion flow both by the external and internal mediums and by the membrane of the cell. It will therefore be appropriately called the *total diffusion resistance of the cell.* It depends both on the physical constants and on the size and shape of the cell.

In the next chapter we shall learn to estimate the value of Λ for some metabolites and shall find it of the order of magnitude 1–10 seconds. This shows that the exponential term in equation (20) will decrease to about one-third of its initial value, which it has for $t = 0$, when $t \sim 1$ to 10 sec. In a next similar interval it will decrease to one-ninth of its initial value, and so on. Unless the initial value of $\bar{c}$ at $t = 0$ is larger by several orders of magnitude than the stationary value as given by equation (21)—a case which is not likely to occur—this stationary value will be practically reached within a minute or less. Thus, with the foregoing value of Λ, a cell kept under constant conditions for a matter of hours will be practically in a stationary state.

It is, however, very important to emphasize that a sta-

tionary state does not always exist. We have considered hitherto only cases in which the rate of production or consumption of a substance by the cell does not depend on the concentration of that substance within the cell. If the rate q is itself a function of $\bar{c}$, so that we may write $q = q(\bar{c})$, then, for the stationary state we still have five algebraic equations, namely, (6)–(9) and (26). But, putting in equation (26) $q = q(\bar{c})$ makes this equation, in general, nonlinear in $\bar{c}$; and under these conditions, for some types of functions $q(\bar{c})$, the whole system of five equations may have no real roots, or the roots may be real but negative, which is physically impossible. This means that under such conditions stationary states do not exist. A study of the general nonstationary case in some such instances reveals that the concentration $\bar{c}$ increases indefinitely with time, regardless of initial conditions. Such is the case when a substance is produced inside the cell at a rate proportional to its concentration, so that $q = k^2\bar{c}$, where k^2 is a constant. It is interesting that in this case a stationary state exists only for cells of sufficiently small size (*MB*, chap. iv). For a rounded-up cell—that is, for $r_1 = r_2 = r_0$—and for $D_e = \infty$ the critical size r^*, above which no stationary state exists, is given[2] by

$$r^* = \frac{D_i}{h}\left(\sqrt{1 + \frac{6h^2}{k^2 D_i}} - 1\right) \tag{27}$$

and by a somewhat more complicated expression for the general case of a finite D_e. However, it is interesting that, when $q = -k^2\bar{c}$—in other words, when the substance is *consumed* at a rate proportional to its own concentration—a stationary state does always exist and is approached regardless of initial conditions.

A study of the foregoing case was made first by solving the differential equations of diffusion exactly for a sphere (*MB*,

chap. iv). Not only is the mathematics somewhat involved, but the expression for r^* could not be obtained explicitly, being given by a transcendental equation. The approximate method shows here also the advantage of both greater simplicity and generality.

It is also possible that, while a stationary state exists, it is not approached asymptotically. The concentration $\bar{c}$ in this case oscillates periodically around the stationary value. For a very special case ($D_i = D_e = \infty$), such a situation was studied by the author (*MB*, chap. vi). Alvin Weinberg[3] worked out subsequently an exact solution for a more general case, namely, $D_e = \infty$, with D_i and h finite. The mathematical procedure is again rather elaborate and lengthy. Such periodical solutions occur when we have at least two substances in the cell, affecting the rates of formation of each other. Weinberg's exact solution has been obtained only for the particular case that the ratio h/D_i is the same for both substances. Using the approximation method, Weinberg[4] finds in a rather simple way expressions of the same form as the exact ones, and without using the foregoing restriction as to the ratio h/D_i. The comparison of the two papers of Weinberg is rather instructive.

For use in subsequent chapters we shall derive, in conclusion, a few auxiliary equations.

From equations (15) and (16) we have

$$c_2 - c_1 = \frac{2(D_iD_e + \delta D_i h)hD_e(r_1 - r_2)(\bar{c} - c_0)}{(2D_iD_e + 2\delta D_i h + r_1 hD_e)(2D_iD_e + 2\delta D_i h + r_2 hD_e)}. \quad (28)$$

Introducing expressions (18) into (21) and the latter into (28), we find

$$\left.\begin{aligned} c_2 - c_1 &= \frac{r_1 r_2 q}{3} \\ &\times \frac{2(D_e + \delta h)(r_1 - r_2)}{2(2D_iD_e + 2\delta D_i h + r_1 hD_e)r_1 + (2D_iD_e + 2\delta D_i h + r_2 hD_e)r_2}. \end{aligned}\right\} (29)$$

For $h = \infty$, equations (28) and (29) reduce, respectively, to

$$c_2 - c_1 = \frac{2\delta D_i D_e}{(2\delta D_i + r_1 D_e)(2\delta D_i + r_2 D_e)} (r_1 - r_2)(\bar{c} - c_0) \quad (30)$$

and

$$c_2 - c_1 = \frac{r_1 r_2 q}{3} \frac{2\delta(r_1 - r_2)}{2(2\delta D_i + r_1 D_e) r_1 + (2\delta D_i + r_2 D_e) r_2}. \quad (31)$$

From equations (6), (7), and (17) we have

$$\left.\begin{aligned} c_1 - c_1' &= \frac{2D_i D_e}{2D_i D_e + 2\delta D_i h + r_1 h D_e} (\bar{c} - c_0) \\ \text{and} \\ c_2 - c_2' &= \frac{2D_i D_e}{2D_i D_e + 2\delta D_i h + r_2 h D_e} (\bar{c} - c_0). \end{aligned}\right\} \quad (32)$$

For $h = \infty$ this gives $c_1 - c_1' = c_2 - c_2' = 0$. Physically, this should be the case, since, when the membrane is infinitely permeable, any finite concentration difference on both sides will result in an infinite flow. Hence, for any finite flow the concentration difference in this case is zero.

REFERENCES

1. Gale Young, *Bull. Math. Biophysics*, 1, 31, 1939. (The *Bulletin of Mathematical Biophysics* is a new quarterly journal devoted to research in that field and containing most of the current publications. Editor: N. Rashevsky. Editorial and publication office: 5822 Drexel Avenue, Chicago, Illinois.)

2. N. Rashevsky, "Note on Autocatalytic Cell Reactions," *Libre de homenagem dos Professores Alvaro e Miguel Ozorio de Almeida* (Rio de Janeiro, 1939).

3. Alvin M. Weinberg, *Growth*, 2, 81, 1938.

4. Alvin M. Weinberg, *Bull. Math. Biophysics*, 1, 19, 1939.

CHAPTER II

COUPLED REACTIONS: APPLICATIONS TO CELL RESPIRATION

We shall now consider a somewhat more complex situation, in which several coupled chemical reactions take place in the cell simultaneously. Actually, a very large number of simultaneous reactions are taking place in any cell, and we cannot hope to tackle the problem in all its complexity. We shall, however, investigate such relatively simple types of reactions as are suggested by actual observations.

In a rather large class of cells we have roughly the following situation.[1,2] The cell consumes some type of sugar from the surrounding liquid, and it also consumes oxygen. The sugar, upon entering the cell, undergoes a series of breakdown processes, during which a number of intermediate products are formed. One of the most conspicuous of these intermediate products is lactic acid, which frequently leaves the cell in appreciable quantities. Within the cell, however, the lactic acid undergoes further changes, as a result of which it combines with oxygen, giving carbon dioxide and water. It is also known that the formation of lactic acid from sugar is partly reversible, a part of the lactic acid being resynthesized into higher sugar.

Following H. D. Landahl,[3] we shall therefore investigate mathematically the following situation: sugar flowing into the cell and there breaking up into lactic acid, part of the lactic acid resynthesizing back into sugar, part leaving the cell, and the rest combining with oxygen to produce carbon dioxide and water, both of which leave the cell. To make our problem defi-

nite, we shall assume that the sugar involved is glucose, $C_6H_{12}O_6$. Furthermore, we shall assume that the breakdown of glucose into lactic acid, according to the equation

$$C_6H_{12}O_6 = 2C_3H_6O_3 , \qquad (1)$$

goes on approximately at a rate proportional to its average concentration, $\bar{c}_1$, in the cell, while the resynthesis of lactic acid is proportional to the square of the latter's concentration, $\bar{c}_2$. As for the rate of oxidation of lactic acid, we shall make the plausible assumption that it is proportional to the product, $\bar{c}_2\bar{c}_3$, of the average concentration $\bar{c}_2$ of lactic acid in the cell and of the average oxygen concentration $\bar{c}_3$ in the cell. The oxidation reaction takes place according to the formula

$$C_3H_6O_3 + 3O_2 = 3CO_2 + 3H_2O . \qquad (2)$$

For each gram-mol of lactic acid, 3 gm-mol of oxygen are required. Hence, 90 gm of lactic acid require for their oxidation 96 gm of oxygen; or, for each gram of oxygen consumed, 0.94 gm of lactic acid is oxidized.

These assumptions do not, of course, represent accurately the actual situation, but they may be used as first approximations. Subsequently we shall discuss other, more complex assumptions, which are closer to the true nature of cellular reactions.

If, as we have done in the preceding chapter, we denote the rate of production of a substance by a positive quantity, then the rates of consumption are negative. Denoting the average rate of consumption of glucose by $-q_1$ gm cm^{-3} sec^{-1}, the average rate of production of lactic acid by q_2 gm cm^{-3} sec^{-1}, and the average rate of consumption of oxygen by $-q_3$ gm

cm^{-3} sec^{-1}; denoting, further, by a, b, and f positive coefficients of proportionality and putting $n = 0.94$, we have[4]

$$q_1 = -a\bar{c}_1 + b\bar{c}_2^2 , \tag{3}$$

$$q_2 = a\bar{c}_1 - b\bar{c}_2^2 - nf\bar{c}_2 c_3 , \tag{4}$$

$$q_3 = -f\bar{c}_2\bar{c}_3 . \tag{5}$$

The amount of carbon dioxide produced per cubic centimeter per second is equal, according to equation (2), to $q_4 = -1.38q_3$ gm, while the amount of water produced per cubic centimeter per second is equal to $q_5 = -0.56q_3$ gm. Remember that q_1 and q_3 are negative, while q_4 and q_5 are positive. The quantity q_2 may be either positive or negative, depending on whether more lactic acid is produced from sugar than is oxidized, or vice versa. In the latter case, if there is a sufficient supply of lactic acid outside, it may flow into the cell, to be oxidized there.

We shall confine ourselves to the three reactions represented by equations (3), (4), and (5). For, when we have calculated the values of $\bar{c}_1$, $\bar{c}_2$, and $\bar{c}_3$ and thus know q_1, q_2, and q_3, we immediately have q_4 and q_5. For carbon dioxide the problem thus reduces to one of a substance produced at a given rate, such as has been treated in chapter i. Water, being incompressible, has practically the same concentration always, $\bar{c}_5 = 1$ gm cm^{-3}; and its rate of production determines its velocity of flow from the cell. It may, at first, appear essential to modify our considerations on diffusion so as to take into account that the flowing water also carries part of the dissolved substances with it. The latter are transported not only by diffusion but also by convection. However, a very general estimate (*MB*, chap. i) shows that, because of the usually very small concentrations of the different metabolites (10^{-3} – 10^{-4} gm cm^{-3}), the transport due to water flow is quite negli-

gible, as compared with the transport due to diffusion, and may be safely neglected.

We shall consider here a stationary diffusion state, assuming that it exists under the conditions specified above. This assumption is borne out by finding that the equations have real positive solutions, as we shall see presently.

Denoting by c_{0l} the external constant concentration of the lth substance, we find, by the same line of argument as used in chapter i, that for each substance the following equation holds in the stationary state:

$$\bar{c}_l = c_{0l} + \Lambda_l q_l \tag{6}$$

The indices l have one of the values 1, 2, 3. . . . . For Λ_l we have expression (18) of chapter i, in which we shall write h_l, D_{il}, and D_{el}, referring to the permeabilities and diffusion coefficients of the different substances. Thus, h_1 denotes the permeability for glucose; h_2, that for lactic acid; etc.

Putting in equation (6) $l = 1, 2$, and 3 and substituting it into equations (3)–(5), we find

$$q_1 = -a(c_{01} + \Lambda_1 q_1) + b(c_{02} + \Lambda_2 q_2)^2 \tag{7}$$

$$q_2 = nq_3 - q_1 , \tag{8}$$

$$q_3 = -f(c_{02} + \Lambda_2 q_2)(c_{03} + \Lambda_3 q_3) . \tag{9}$$

Putting

$$\left.\begin{aligned} A &= b\Lambda_2^2 , & H &= f\Lambda_2\Lambda_3 , \\ B &= 2bc_{02}\Lambda_2 , & K &= 1 + fc_{02}\Lambda_3 , \\ C &= a\Lambda_1 + 1 , & M &= fc_{03}\Lambda_2 = M_0 c_{03} , \\ D &= ac_{01} - bc_{02}^2 , & S &= fc_{02}c_{03} = S_0 c_{03} , \\ \Lambda_0 &= f\Lambda_2 , & S_0 &= fc_{02} , \end{aligned}\right\} \tag{10}$$

we may write equations (7), (8), and (9) thus:

$$Aq_2^2 + Bq_2 - Cq_1 - D = 0\,, \tag{11}$$

$$q_1 + q_2 - nq_3 = 0\,, \tag{12}$$

$$Hq_2q_3 + Kq_3 + Mq_2 + S = 0\,. \tag{13}$$

These are three equations for the three unknown quantities q_1, q_2, and q_3. If we solve these equations, we can immediately find $\bar{c}_1$, $\bar{c}_2$, and $\bar{c}_3$ by using equation (6). But equations (11), (12), and (13) are, except for the different convention regarding the signs of the q's, the same as equations (22), (23), and (24) of chapter v of *MB*, obtained previously by H. D. Landahl by a different method, and their solution is the same as given there. While, however, the equations discussed in *MB* held only for spherical cells, the equations given here and obtained by the general approximation method hold for cells of any oblong shape.

Equations (11), (12), and (13) determine q_1, q_2, and q_3 in terms of such constants as the diffusion coefficients, permeabilities, and the external concentrations c_{01}, c_{02}, and c_{03}. We shall confine ourselves here to the study of the average oxygen consumption q_3 in its dependence on the external oxygen concentration c_{03}, as there are a number of experimental data with which our conclusions in this respect may be compared.

Solving equations (11), (12), and (13) for q_3, we find that for $c_{03} = \infty$, q_3 tends to a constant limiting value, q_3^*, given by (*MB*, p. 48):

$$q_3^* = \frac{(a\Lambda_1 + 1)c_{02} + a\Lambda_2 c_{01}}{n\Lambda_2(a\Lambda_1 + 1)} \tag{14}$$

Considering, instead of q_3, the quantity $y = q_3/q_3^*$ (that is, the relative oxygen consumption) and putting $c_{03} = x$, we find from equations (11), (12), and (13) that, approximately,

$$x = \zeta y + \frac{\xi y}{1 - y} \tag{15}$$

where

$$\xi = \frac{\perp}{nf\Lambda_2}; \qquad \zeta = -\Lambda_3 q_3^* \tag{16}$$

For $y = 1$, $x = \infty$, while for $y = 0$, $x = 0$. That is, if we plot y against x, we obtain a curve rising from zero and approaching asymptotically the constant value 1. Such should be the relation between the relative oxygen consumption y and the external oxygen concentration x, all other conditions

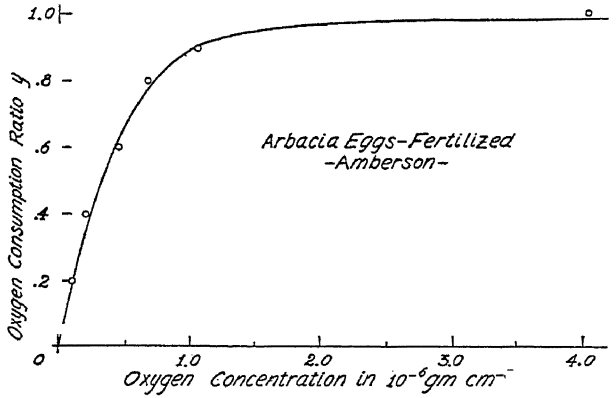


FIG. 2.—The curve represents the theoretical equation (15). The points represent experimental data.[5] $\Lambda_3 = 2$ sec; $D_3 \gtrsim 8 \times 10^{-7}$ cm^2 sec^{-1}; $h_3 \gtrsim 4.2 \times 10^{-4}$ cm sec^{-1}.

being kept constant. To what extent equation (15) is verified experimentally is shown in Figures 2–7.

Inasmuch as the approximation method used here applies to any shape of cell, it should apply also to pieces of tissue consisting of a large number of cells. The diffusion coefficient for such a piece of tissue is, of course, only a sort of average quantity, but the same holds for any cell that is physically heterogeneous. Figure 7 illustrates how the equations here developed apply to tissues.

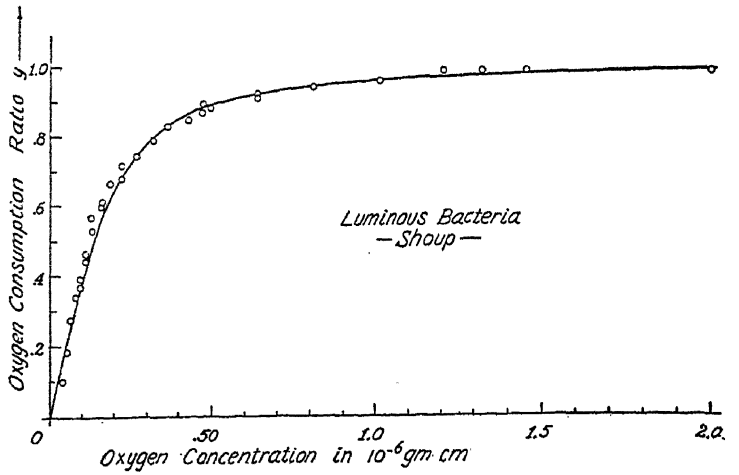


FIG. 3.—The curve represents the theoretical equation (15). The points represent the experimental data.[6] $\Lambda_3 = 4.3 \times 10^{-4}$ sec; $D_3 \geq 1.3 \times 10^{-6}$ cm^2 sec^{-1}; $h_3 \geq 3.6 \times 10^{-2}$ cm sec^{-1}.

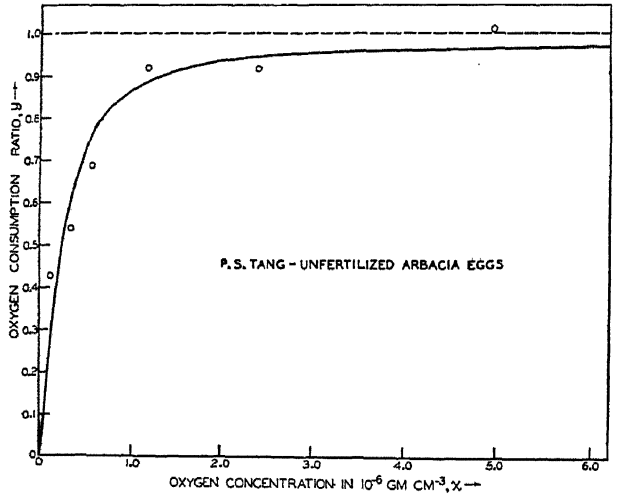


FIG. 4.—The curve represents the theoretical equation (15). The points represent experimental data.[7] $\Lambda_3 = 4$ sec; $D_3 \geq 4 \times 10^{-7}$ cm^2 sec^{-1}; $h_3 \geq 2 \times 10^{-4}$ cm sec^{-1}.

In some of the experiments, namely, those with *Arbacia* eggs, the conditions were, however, not such as assumed in the derivation of equation (15). In the experiments with *Arbacia* eggs the latters were kept in saline solutions, practically

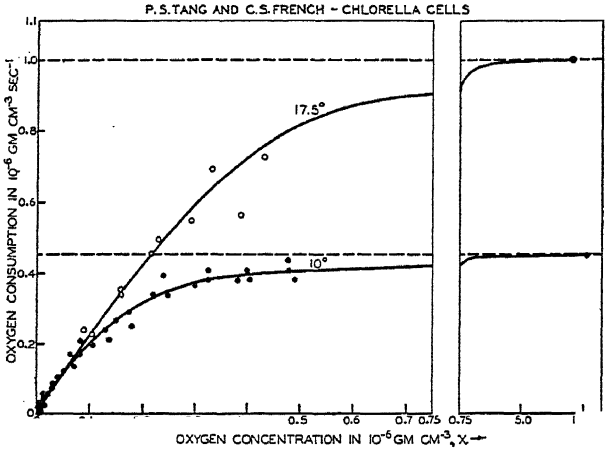


FIG. 5.—The curves represent the theoretical equation (15). The points represent experimental data.[8] Observations were made at different temperatures, as indicated on the graphs (degrees centigrade). A different scale for x is used for $x > 0.75 \times 10^{-6}$ gm cm^{-3}, in order to represent the whole range of data in the same figure. The constants calculated for the two different temperatures are as follows: For 10° C: $\xi = 4 \times 10^{-8}$ gm cm^{-3}; $\zeta = 1.5 \times 10^{-7}$ gm cm^{-3}; $\Lambda_3 = 0.33$ sec; $D_3 \geq 1.4 \times 10^{-8}$ cm^2 sec^{-1}; $h_3 \geq 1.4 \times 10^{-4}$ cm sec^{-1}. For 17°.5 C: $\xi = 3.7 \times 10^{-8}$ gm cm^{-3}; $\zeta = 4.2 \times 10^{-7}$ gm cm^{-3}; $\Lambda_3 = 0.42$ sec; $D_3 \geq 1.1 \times 10^{-8}$ cm^2 sec^{-1}; $h_3 \geq 1.1 \times 10^{-4}$ cm sec^{-1}.

free from sugar. The eggs in this case oxidize their own reserve of sugar stored within the cell. This reserve is apparently large enough to maintain for a sufficient time a constant supply and thus secure a quasi-stationary state. The theory for this case requires, however, only a very slight modification. All we have to do[3] is to substitute in equations (3) and

(4) a constant, a_1, for the term $a\bar{c}_1$. That constant, a_1, represents the constant rate of decomposition of the sugar reserve. However, this modification of the equations does not alter the final equation (15) (*MB*, chap. v).

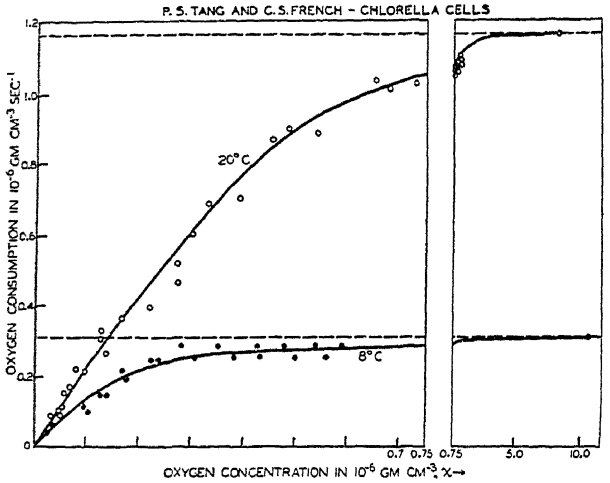


FIG. 6.—Same as Fig. 5, with constants as follows: For 8° C: $\xi = 4 \times 10^{-8}$ gm cm^{-3}; $\zeta = 1.5 \times 10^{-7}$ gm cm^{-3}; $\Lambda_3 = 0.48$ sec; $D_3 \geq 1 \times 10^{-8}$ cm^2 sec^{-1}; $h_3 \geq 0.9 \times 10^{-4}$ cm sec^{-1}. For 20° C: $\xi = 3.3 \times 10^{-8}$ gm cm^{-3}; $\zeta = 5 \times 10^{-7}$ gm cm^{-3}; $\Lambda_3 = 0.43$ sec; $D_3 \geq 1.1 \times 10^{-8}$ cm^2 sec^{-1}; $h_3 \pm 1.1 \times 10^{-4}$ cm sec^{-1}.

From the comparison of the experimental data with equation (15) we obtain numerical values for ζ and ξ. Since in the experiment the quantity q_3^* is given as the absolute maximum rate of oxygen consumption per cubic centimeter, knowledge of ξ gives us the value of Λ_3, the total diffusion resistance for oxygen. The value of Λ_2, unfortunately, cannot be obtained from the knowledge of ζ unless we know the coefficient f.

In experiments made with *Arbacia* eggs and with *Chlorella*

cells the suspension of the latter was thoroughly stirred in a respirometer. Such stirring uniformizes the concentrations of substances outside the cell and is equivalent in its effects to making D_e very large. In the expression for Λ_3 we may there-

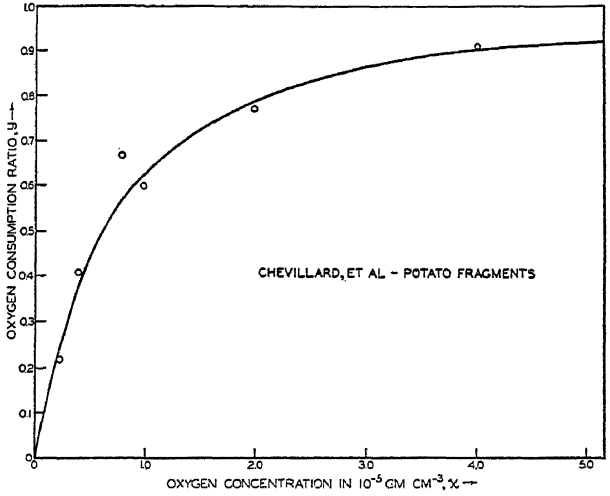


FIG. 7.—Comparison of the theoretical equation (15), as represented by the curve, with experimental data, taken from a paper by P. S. Tang.[9] Since the size of the fragments were not given, the constant Λ_3 could not be computed.

fore put, with a good approximation, $D_e = \infty$. Moreover, since the *Arbacia* eggs are spherical, we have $r_1 = r_2 = r_0$. In this case the expression (18) of chapter i gives

$$\Lambda_3 = \frac{2}{9}\frac{r_0}{h_3} + \frac{r_0^2}{9D_{i3}} \qquad (17)$$

While r_0 is given experimentally, the knowledge of Λ_3 does not yet give us either h_3 or D_{i3} separately. Since, however, both terms on the right-hand side of equation (17) are posi-

tive, therefore at least one of them, or both, must be of the order of magnitude of Λ_3. Hence, we have

$$D_{i_3} \gtrsim \frac{r_0^2}{9\Lambda_3} \quad \text{and} \quad h_3 \gtrsim \frac{2r_0}{9\Lambda_3}, \tag{18}$$

where at least one or both of the equivalence ($\sim$) signs must hold. That is, if $D_{i_3} \gg r_0^2/9\Lambda_3$, then certainly $h_3 \sim 2r_0/9\Lambda_3$; and, conversely, if $h_3 \gg 2r_0/9\Lambda_3$, then $D_{i_3} \sim r_0^2/9\Lambda_3$. But both equivalence signs may hold simultaneously.

Values of the different constants estimated in this way are given in the legends accompanying Figures 2–6.

In principle h_3 and D_{i_3} could be obtained by experimenting on two different groups of cells of the same type but of different size. This would supply us with two equations of the type of equation (17), one for each value of r_0. The objection, however, could be raised that cells of different size may have different values of h_3 and D_{i_3}. The equations developed here may, however, be applied not only to single cells but to slices of tissues of different size and shape. The expression for Λ_3 will, of course, now be of a more complex structure, though we could still have $D_e = \infty$. One might hope to separate, in this way, h_3 from D_{i_3}. However, in the case of a multicellular slice the inner meshwork of the membranes would add to the diffusion resistance due to D_{i_3}, and we would have to modify somewhat the equations to take this into account.

Data by P. S. Tang and C. S. French[8] represented on Figures 5 and 6 were taken for different temperatures. Determining for each temperature the parameters ζ and ξ, we may study their variation with temperature, as is shown on Figure 8. As we see, the quantity ζ increases with temperature. The exact variation cannot be ascertained from only four points. But, since ζ is proportional to the rate of reaction q_3^* (cf. Eqs. [16]), we would expect ζ to increase with tempera-

ture, since all chemical reactions do so. For ξ we should expect a decrease, since the reaction constant f would be expected to increase.

In many cell-kinetic problems, however, only the total diffusion resistance Λ of the cell is of consequence, and it is therefore important to be able to determine these quantities, even if we cannot resolve them into their components.

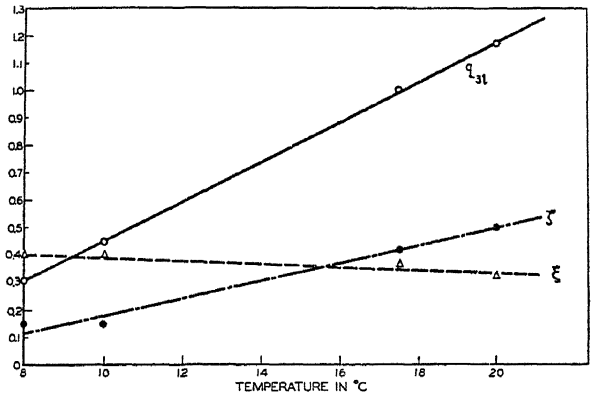


FIG. 8.—Variation of the quantities q_{3l} (full line), ξ (broken line), and ζ (alternate line) with temperature, as computed from Figs. 5 and 6. The scale for q_{3l} is in 10^{-6} gm cm^{-3} sec^{-1}; that for ξ, in 10^{-7} gm cm^{-3}; and that for ζ, in 10^{-6} gm cm^{-3}.

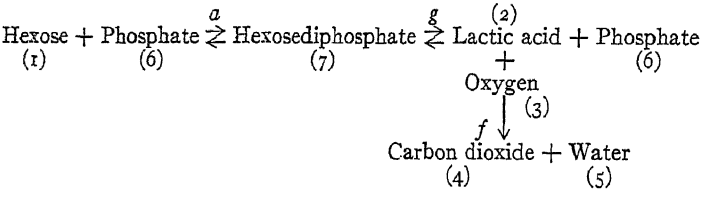
The quantity q_3^* is, as seen from equation (14), a function of the external concentration c_{02} of lactic acid and of the concentration c_{01} of glucose. Unfortunately, there are no data available to check this relation. Should, however, such data be obtained and should equation (14) be found to be sufficiently accurate, this would immediately give us the values of Λ_1 and Λ_2, the total diffusion resistances for glucose and lactic acid. Other parameters may be determined by studying the consumption of glucose and the output of lactic acid in their dependence on the external concentration of glucose, lactic

acid, and oxygen. The urgent need of further experimental work on cell respiration under conditions just mentioned is thus clearly indicated.

If we deal with cells having a very thin membrane, such that its physical constitution is not likely to be very different from that of the rest of the cytoplasm, then the contribution of such a membrane to the total diffusion resistance is negligible, and we may, with a good approximation, make in equation (17) $h = \infty$, which gives us D_{i_3} in terms of Λ_3. In many experimental determinations of cell permeability, especially in those using plasmolysis of plant cells, it is the permeability of the relatively thick layer of cytoplasm plus surface membrane that is determined. What we actually get there is, therefore, a quantity proportional to $1/\Lambda$.

H. D. Landahl has also studied more complex situations. For instance, we may consider[4] the case where the resynthesis of sugar occurs in two routes—one direct and proportional to $b\bar{c}_2^2$, and the other coupled to the energy liberated by the oxidation of part of the lactic acid and to the rate $f\bar{c}_2\bar{c}_3$ of oxygen consumption. As far as the dependence of oxygen consumption on oxygen pressure is concerned, we obtain equation (15).

Some interesting results are obtained by considering the following type of reaction.[4]



In this scheme the number under or above each metabolite shows the index which refers to it in the following equations. Thus, $\bar{c}_7$ means the average concentration of hexosediphosphate; q_6, the average rate of production per cubic centimeter

of phosphate, etc. The letters a, g, and f next to the arrows are the coefficients of proportionality of the corresponding reactions, used in the equations which follow. Let us assume that the phosphate never leaves the cell, in any form, and that it neither accumulates nor is depleted, so that the total amount of phosphate per cubic centimeter is a constant, c_p. For simplicity we shall consider the reactions as not reversible. Also, again the equations for carbon dioxide and water shall be omitted, for the same reasons as before. Denoting by M_l the molecular weight of the lth substance and putting $m = 2M_6/M_7$ and $n = 2M_2/6M_3$, we have for this case, instead of equations (3)–(5), the following set of equations:

$$q_1 = -a\bar{c}_1\bar{c}_6 , \tag{19}$$

$$q_2 = -nf\bar{c}_2\bar{c}_3 + g\bar{c}_7 , \tag{20}$$

$$q_3 = -f\bar{c}_2\bar{c}_3 , \tag{21}$$

$$q_6 = -q_7 = -a\bar{c}_1\bar{c}_6 + g\bar{c}_7 = 0 , \tag{22}$$

$$c_p = \bar{c}_6 + m\bar{c}_7 . \tag{23}$$

Consider first the anaerobic case, that is, the case when there is no external supply of oxygen, so that $c_{03} = \bar{c}_3 = q_3 = 0$. Then, solving equation (23) for $\bar{c}_7$, substituting into equation (22), and rearranging, we find

$$\bar{c}_6 = \frac{gc_p}{am\bar{c}_1 + g} . \tag{24}$$

Substituting into equation (19) for $\bar{c}_6$ the value given by equation (24) and remembering that $\bar{c}_1 = c_{01} + \Lambda_1 q_1$, we obtain

$$am\Lambda_1 q_1^2 + (amc_{01} + agc_p\Lambda_1 + g)q_1 + agc_{01}c_p = 0 . \tag{25}$$

Putting

$$q_1^* = -\frac{gc_p}{m}; \qquad z = \frac{q_1}{q_1^*}; \qquad \zeta' = \frac{gc_p\Lambda_1}{m}; \qquad \xi' = \frac{\xi}{am}, \tag{26}$$

we may write equation (25) thus:

$$c_{01} = \zeta' z + \frac{\xi' z}{1 - }. \tag{27}$$

Equation (27) gives a relation between the external concentration of sugar and the rate of sugar consumption. In absence of oxygen the rate of output of lactic acid is proportional to the rate of sugar consumed. Therefore equation (27) also gives us a relation between rate of lactic-acid output and sugar concentration. Equation (27) is of the same form as equation (15), showing that for zero sugar concentration the lactic-acid output is zero and that the latter tends to a constant limiting value for large sugar concentrations.

Data on this point are very meager; still, the following considerations may be of some interest. On Figure 9 data are plotted for lactic-acid output from experiments with glucose and fructose. If we use different scales for the abscissae, then the two sets of points *seem* to fall on the same curve, represented by equation (27). Whether this is so or not, cannot be decided, because of a too small number of points. But assuming, for a moment, that this is so, we are led to the following rather interesting considerations. If the curves did coincide in the figure, then $\xi'_{fr} = 20\xi'_{gl}$ and $\zeta'_{fr} = 20\zeta'_{gl}$, where the subscripts refer to fructose and glucose and where the factor 20 arises because there happens to be a factor of 20 between the two scales. Since the limiting values of the lactic-acid production is nearly the same in each,[10] then by equation (27) we have $\Lambda_{1_{fr}} = 20\Lambda_{1_{gl}}$, or the total diffusion resistance for fructose is twenty times that for glucose. If we assume, also, that

the secondary stage is, to a large extent, the same in both cases, so that c_p is nearly the same in each, then, since m is the same for both, g must be the same for both to satisfy equation (27). In other words, the second-stage reaction goes on at about the same rate in both. In this case one might

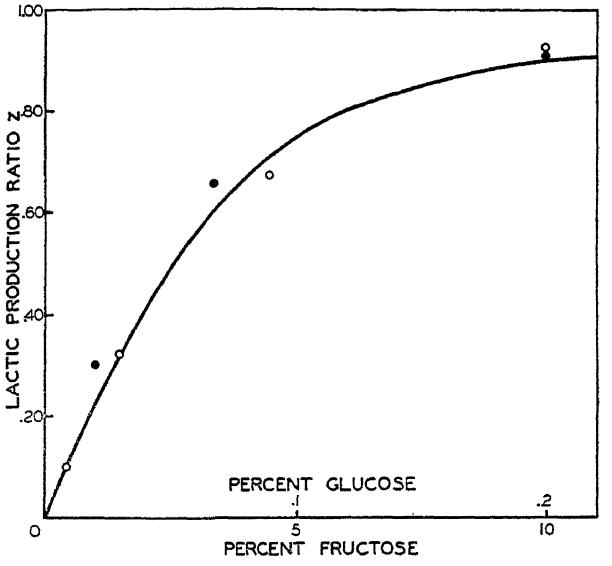


FIG. 9.—The curve represents equation (27). The points represent experimental data. For explanation of double scale on the axis of abscissae see text, p. 32.

then tentatively assume the secondary reactions to be largely equivalent. From equation (26) we have $a_{gl} = 20a_{fr}$, so that the rate of formation of the second product, hexosediphosphate, from glucose is twenty times the rate of its formation from fructose. However, it could very well be that, at least in some instances, the two curves could not possibly be made to coincide by a linear change in the abscissa scale. In this case it would not be possible to compare the total diffusion

resistance or other constants directly, but it would be necessary to fit the curves empirically to evaluate ξ' and ζ'. The quantity Λ_1 would then be given; but, in general, an additional empirical relation would be required to evaluate a, g, and c_p.

These considerations indicate the importance of accurate determinations of anaerobic lactic-acid production in its dependence on the concentration of different sugars.

For the dependence of oxygen consumption on oxygen pressure the present case leads again to equation (15). The same equation is again obtained by considering a still more complex system,[4] involving the presence of a catalyst necessary for the oxidation of lactic acid. All this shows that the oxygen-consumption–oxygen-pressure curves are determined by the gross features of the phenomena and do not depend much on intermediate details. The reason for the curves being such as shown in Figures 2–7 is physically this: If the concentration of lactic acid in the cell were constant, then increasing indefinitely the oxygen concentration would increase also the oxygen consumption indefinitely. But, as the oxygen concentration, and hence the oxygen consumption, increases, so also does the consumption of lactic acid increase, and therefore the concentration of the latter in the cell decreases. Since the oxygen consumption is proportional to the product of the concentrations of lactic acid and of oxygen, therefore with increasing oxygen concentration and decreasing lactic-acid concentration the oxygen consumption increases less and less rapidly, approaching a limiting value, which is determined by the maximum output of lactic acid due to the breakdown of glucose.

In the next chapter we shall use an important constant of the cell, which we call the "glycolytic coefficient," β. This coefficient β is defined in the following way: 3β is the ratio of molecules of sugar glycolyzed into lactic acid to those oxidized completely. If 1 mole-

cule of glucose is oxidized completely, then 3β are glycolyzed; and altogether $3\beta + 1$ molecules of sugar are consumed. For the oxidation of 1 molecule of sugar, 6 molecules of oxygen are necessary, while 6β molecules of lactic acid and 6 molecules of carbon dioxide are produced. Hence, β is equal to the ratio of the number of molecules of lactic acid produced to the number of molecules of oxygen consumed. In other words,

$$\beta = \frac{-q_2}{M_2}\frac{M_3}{q_3} = -\frac{q_2}{3nq_3} \quad ; \qquad n = \frac{M_2}{3M_3} = 0.94 . \tag{28}$$

Since $q_3 < 0$, β is positive when $q_2 > 0$, that is, lactic acid leaves the cell.

Solving equation (13) for q_2, using notation (10), and substituting the value so obtained into equation (28), we find

$$\beta = \frac{q_3 + fc_{02}(c_{03} + \Lambda_3 q_3)}{3nf\Lambda_2 q_3(c_{03} + \Lambda_3 q_3)} . \tag{29}$$

Putting, again, $y = q_3/q_3^*$ and using expression (16) for ζ and ξ, we have, from equation (15),

$$c_{03} + \Lambda_3 q_3 = c_{03} - \zeta y = \frac{\xi y}{1 - y} . \tag{30}$$

Introducing equation (30) into equation (29) and using equations (16) and (14), we find

$$\beta = \frac{(a\Lambda_1 + 1)}{3[(a\Lambda_1 + 1)c_{02} + ac_{01}\Lambda_2]} \left(\frac{ac_{01}\Lambda_2}{a\Lambda_1 + 1} \frac{1 - y}{y} - c_{02} \right) , \tag{31}$$

which gives us the glycolytic coefficient in terms of c_{01}, c_{02}, and c_{03} (through y, by means of equation [15]). For small values of external lactic-acid concentration c_{02}, equation (31) reduces to

$$\beta = \frac{1}{3} \frac{1 - y}{y} , \tag{32}$$

and for large values of external oxygen concentration c_{03} and therefore $y \backsim 1$, equation (32), because of equation (30), reduces to

$$\beta = \frac{1}{3nf\Lambda_2 c_{03}} . \tag{33}$$

For sufficiently large values of external oxygen concentration c_{03}, the glycolytic coefficient β approaches zero, and may even become negative if c_{02} is sufficiently large. The glycolytic coefficient is inversely proportional to c_{03}. For a constant and not too high value of c_{03}, β increases with decreasing total diffusion resistance, Λ_2; and to a lesser extent β increases as the external glucose concentration, c_{01}, increases; as the total diffusion resistance to oxygen, Λ_3, increases, and as the total diffusion resistance to glucose, Λ_1, decreases. Also, β decreases with f and, to a lesser extent, increases with a.

REFERENCES

1. W. M. Bayliss, *Principles of General Physiology* (4th ed.), chap. ix (New York: Longmans, Green & Co., 1927).

2. *Handbuch d. norm. u. pathologischen Physiologie*, 1, 29 (Berlin: J. Springer, 1927).

3. H. D. Landahl, *Growth*, 1, 263, 1937. Notice the difference in the adapted convention for the sign of q here and in Landahl's paper.

4. H. D. Landahl, *Bull. Math. Biophysics*, 1, 1, 1939. Cf. preceding reference.

5. W. Amberson, *Biol. Bull.*, 55, 73, 1928; also *Tabulae biologicae periodicae*, 3, 52, 1933.

6. C. Shoup, *Jour. Gen. Physiol.*, 13, 27, 1929.

7. P. S. Tang, *Biol. Bull.*, 60, 242, 1931.

8. P. S. Tang and C. S. French, *Chinese Jour. Physiol.*, 7, 353, 1933.

9. P. S. Tang, *Quart. Rev. Biol.*, 8, 260, 1933.

10. E. Boyland and M. Boyland, *Biochem. Jour.*, 32, 321, 1938.

CHAPTER III

DIFFUSION FORCES AS A POSSIBLE CAUSE OF CELL DIVISION

As we have remarked in chapter i, a flow of diffusing substance must exert a force on the medium in which the diffusion takes place, the direction of the force being the same as the direction of flow. The exact theory of these forces presents rather great difficulties because it involves a detailed knowledge of the molecular structure of liquids, especially of colloids and gels, in which the diffusion usually takes place. A pure solvent, like water, may behave differently, in this respect, from a complex gel. In order not to lose ourselves in generalities, it is well to make at once, at least in a preliminary way, some rough picture of the structure of the inside of the cell. We shall first adopt the simplest possible scheme, later on discussing possible complications.

Obviously, if the cell consisted merely of a sac, formed by the membrane and containing only the same solvent and solutes as are present outside, such a cell would not exhibit the simplest fundamental properties of actual cells. While, for instance, sugar and oxygen are present in sufficient concentrations, no appreciable oxidation occurs outside the cell. Other important reactions also take place only inside the cell, though the necessary components may be present outside. This must be due to the presence inside the cell of various catalysts, to which the cell membrane is not permeable, so that they are confined within the cell. Most biological catalysts, or enzymes, are known to be attached to various colloidal structures.[1] We may picture, therefore, the inside of the cell as being composed of water, in which are suspended small col-

loidal particles. Water is the common medium outside and inside the cell. The cell proper consists, then, according to this picture, of an aggregate of colloidal particles, inclosed in a membrane.

The drag forces, exerted by the flow of diffusing substances, will be acting both on the water and on the particles. We are interested especially in the forces acting on the particles, which represent, in a way, the "skeleton" of the cell.

Considering that the diffusion flow always occurs as a result of nonuniformities of concentrations, and that in a field of nonuniform concentration of a solute any particle, as well as any molecule, of the solvent is subject to an asymmetric bombardment by molecular collision of the nonuniformly distributed diffusing solute, the author derived (*MB*, chap. vii) the following expression for the force F exerted on a spherical particle of volume V_p:

$$F = -\frac{3}{2}\frac{RT}{M}\alpha V_p \operatorname{grad} c\,. \qquad (1)$$

Here R denotes the gas constant per mol ($R = 0.83 \times 10^8$ ergs degree^{-1}); T, the absolute temperature; M, the molecular weight of the diffusing substance; and α, a constant, less than unity but of the order of unity, and determined by the molecular structure of the solvent. The symbol grad c (read "gradient of c") means the change of the concentration c of the diffusing substance per unit length, in the direction of the greatest change. The minus sign (−) shows that the force is directed from higher to smaller concentrations, that is, in the direction of flow.

Gale Young[2] pointed out certain inadequacies in the author's derivation of equation (1). He derives a similar equation by considering, instead of the molecular bombardment, the viscous drag exerted by the flowing substance upon a

spherical particle. He obtains a formula very similar to equation (1) except that the quantity $-3RTV/2M$ is not multiplied by a correction factor a but has a very small additive correction term. Inasmuch as both equations are derived under rather simplified assumptions, they cannot be expected to give more than a close order of magnitude; and no choice is indicated, at present, between them. It is, however, important that two entirely different approaches lead to essentially the same expressions.

If we have N_p particles per unit volume, then the force acting on a cubic centimeter of the solid cell constituents is approximately equal to

$$f = -\frac{3}{2}\frac{RT}{M} N_p V_p \operatorname{grad} c . \qquad (2)$$

But $N_p V_p$ is nothing but the actual volume occupied by the particles contained in a cubic centimeter, a quantity which we shall denote by μ. We thus have for the force acting on the cell constituents per unit volume

$$f = -\frac{3}{2}\frac{RT\mu}{M} \operatorname{grad} c . \qquad (3)$$

Actually the colloidal particles constituting the body of the cell may not be spherical. From general considerations it appears plausible that for any shape of particles an equation similar to equation (1) will hold approximately, except for a different numerical coefficient[2] (also *MB*, chap. vii). We may now imagine the particles to be held together by highly viscous or even solid strands, thus forming a quasi-solid meshwork through which water and substances dissolved in the latter circulate freely. This conforms in general with accepted ideas on the structure of some jellies. In that case an equation of

the form of (2) will still represent the force exerted by the diffusion flow on the unit volume of the cell. The coefficient μ now denotes the relative true volume occupied by the solid structure. When we shall speak of the deformation of the cell, we shall refer to the deformation of this structure, consisting either of a jelly-like rigid network or of disconnected solid particles. Water itself, which may either be stationary or circulate, is considered just as a special metabolite.

We may now proceed to calculate the mechanical effects of such forces on the cell. We must remember that each metabolite produces a force of its own and that the total force is the sum of the individual components due to each metabolite. Let us first investigate the effect of a single substance and then pass to the more complex case, which will enable us to compare with experimental data the equations derived.

Let us again consider a substance produced in the cell at a rate q gm cm^{-3} sec^{-1}. The flow in this case is directed outward, and the forces are also directed outward. They have, therefore, the tendency to disrupt the cell. For a case of a perfect sphere, due to the complete symmetry of everything with respect to the center, the resultant force will be zero. But any deviation from the spherical shape will result in an asymmetry in the distribution of the forces; the resultant effect will not be zero, and a deformation of some sort will occur (*MB*, chap. ix).

In the case of a cell producing a substance we, first of all, have a force, given by equation (3), acting on each element of volume. Moreover, at the membrane of the cell there is a pressure acting on the membrane because of the difference in the concentrations on both sides of the membrane. This difference in concentrations results in a difference of osmotic pressures. The osmotic pressure of a solution of concentration c is equal to

$$p = \frac{RT}{M} c \text{ dynes cm}^{-2}. \tag{4}$$

The net pressure on the membrane is therefore the difference of the pressure due to the inside concentration, directed outward, and of the pressure due to the outside concentration, directed inward. For the "ends" and the "sides" of the cell (chap. i, Fig. 1) this resulting pressure is equal, respectively, to

$$\begin{aligned} p_1 &= \frac{RT}{M}(c_1 - c_1') \text{ dynes cm}^{-2}; \\ p_2 &= \frac{RT}{M}(c_2 - c_2') \text{ dynes cm}^{-2}. \end{aligned} \tag{5}$$

Whether we regard the protoplasm constituting the cell as a sol or as a semisolid gel, the cell must be considered as a body possessing an appreciable amount of viscosity. Any deformation resulting in relative displacement of the different parts of the cell with respect to each other will set up viscous resistance forces which will be stronger the higher the rate of deformation.

In order to calculate the deformation of the cell under the influence of the forces discussed above, we shall make use of a theorem by Betti,[3] combined with the laws of plastic flow.[4] If l_z is the length of the body in a given direction (say in the direction of the z-axis), and if V is the volume of the body, S its surface, η its viscosity, X, Y, and Z the components of the volume force per unit volume in the direction of the corresponding coordinate axes, and X_ν, Y_ν, and Z_ν the components of pressure at the surface, then, according to Betti's theorem, the average relative rate of change $(1/l_z)(dl_z/dt)$ of l_z is given, for a body of *any* shape, by[4]

$$\left.\begin{aligned} \frac{1}{l}\frac{dl_z}{dt} = \frac{1}{3\eta V}\Big\{ &\iiint_V [zZ - \tfrac{1}{2}(yY + xX)]dV \\ &+ \iint_S [zZ_\nu - \tfrac{1}{2}(yY_\nu + xX_\nu)]dS \Big\}; \end{aligned}\right\} \tag{6}$$

where the first integral is extended over the whole volume, and the second over the whole surface of the body.

Equation (3) shows that the volume force f can be derived from a potential $\varphi = \frac{3}{2}(RT\mu/M)c$. In this case, as has been shown by Gale Young,[4] the volume integral in equation (6) can be easily transformed into a surface integral of the form

$$-\frac{3RT\mu}{2M}\iint_S c\{z\cos(\nu, z) - \tfrac{1}{2}[x\cos(\nu, x) + y\cos(\nu, y)]\}dS\,.$$

Equations similar to (6) hold for all three directions. If we find that the body elongates in two directions—that is, if for two directions the right-hand side of equation (6) is positive—then we shall find that it contracts in the third direction, the right-hand side of equation (6) being, for this direction, negative. If the body possesses approximately axial symmetry, as we assume to be the case for a cell of the type considered in chapter i, then elongation in one direction results in a constriction in the other two directions.

Putting the z-axis in the direction of the largest dimension of the cell, we have $l_z = r_1$. Putting

$$I_1 = \iint_S c\{z\cos(\nu, z) - \tfrac{1}{2}[x\cos(\nu, x) + y\cos(\nu, y)]\}dS \quad (7)$$

and

$$I_2 = \iint_S [zZ_\nu - \tfrac{1}{2}(xX_\nu + yY_\nu)]dS\,, \quad (8)$$

where X_ν, Y_ν, and Z_ν are the components of the pressure given by equation (5), and using the above-mentioned transformation of the volume integral, we have

$$\frac{1}{r_1}\frac{dr_1}{dt} = -\frac{RT\mu}{2M\eta V}I_1 + \frac{1}{3\eta V}I_2\,. \quad (9)$$

We now proceed to calculate the quantities I_1 and I_2. They require a knowledge of the concentration at the surface of the cell. Strictly speaking, when the shape of the cell changes with respect to time, we cannot take for the concentration inside the cell at any given moment the values obtained for the stationary state of the cell having the same shape as the given cell has at this moment, for the latter were obtained in chapter i under the assumption that the shape of the cell does not vary. However, we have seen that, beginning with any initial state, the stationary state is practically reached within a couple of minutes. If the deformation of the cell proceeds very slowly, so that its shape does not change much within those few minutes, then we can apply with sufficient accuracy the expression derived for the stationary state. In many cases the time it takes a cell to divide is of the order of magnitude of hours. The time necessary to readjust the diffusion state to a different shape being much shorter, the system may be considered as a succession of stationary states. However, the more general case, when the rates of deformation are not slow enough, must also be studied. This will be done in chapter v.

In calculating I_1 we remark that on the "sides" of the cell $\cos(\nu, z) = 0$, while $x \cos(\nu, x) + y \cos(\nu, y) = r_2$. Remembering, from chapter i, that the volume V of the cell is equal to

$$V = \tfrac{4}{3}\pi r_1 r_2^2 , \tag{10}$$

we find as a contribution to I_1 for the "sides"

$$-c_2 \times \tfrac{1}{2}r_2 \times 2\pi r_2 \times 2r_1 = -\tfrac{3}{2}Vc_2 . \tag{11}$$

In a similar way we find that the contribution of the "ends" to I_1 is $\frac{3}{2}Vc_1$. Hence,

$$I_1 = \tfrac{3}{2}V(c_1 - c_2) . \tag{12}$$

In calculating I_2 we have along the "sides"

$$xX_\nu + yY_\nu = \frac{RT}{M} r_2(c_2 - c_2') ,$$

while at the "ends"

$$zZ_\nu = \frac{RT}{M} r_1(c_1 - c_1') .$$

Therefore,

$$I_2 = \tfrac{3}{2} V \frac{RT}{M} [(c_1 - c_1') - (c_2 - c_2')] \qquad (13)$$

Introducing equations (12) and (13) into equation (9) and introducing into that the expressions (28) and (32) of chapter i, we find

$$\frac{1}{r_1}\frac{dr_1}{dt} = \frac{RT}{2M\eta} \times \frac{[3\mu\delta h + (3\mu - 2)D_e]hD_iD_e(r_1 - r_2)(\bar{c} - c_0)}{(2D_iD_e + 2\delta D_i h + r_1 hD_e)(2D_iD_e + 2\delta D_i h + r_2 hD_e)} . \qquad (14)$$

For $r_1 = r_2$ the deformation dr_1/dt vanishes, as should be the case.

For $q > 0$, that is, when the substance is produced, $\bar{c} - c_0 > 0$. The quantity $(3\mu - 2)$ is usually negative. Since $r_1 - r_2 > 0$, therefore for sufficiently small $\delta \sim r_1 \sim r_2$, $dr_1/dt < 0$ for $q > 0$. If, however, the quantities $\delta \sim r_1 \sim r_2$ are sufficiently large, then the numerator becomes positive, and $dr_1/dt > 0$ for $q > 0$. In other words, omitting the rather unlikely case that $\mu > \frac{2}{3}$, for cells of a sufficiently small size the effect of the diffusion forces generated by a produced substance is such that any oblong shape tends to round up. But when the size of the cell is sufficiently large, the effect of such

forces is to elongate the cell. Hence, a substance produced by the cell will always cause an elongation of the cell as soon as the latter grows sufficiently large. To insure positivity of dr_1/dt we must have

$$\delta > \frac{2 - 3\mu}{3\mu} \frac{D_e}{h} \sim 10^{-3} \text{ cm}. \tag{15}$$

If we consider a consumed substance, we have $q < 0$, and the situation is reversed. The diffusion forces tend to elongate the cell only when it is small enough. As we shall presently see, the surface tension of the cell always opposes the elongation. But the effect of surface tension decreases with increasing size of the cell. Therefore, for $q > 0$, that is, for a produced substance, for large sizes the diffusion forces will prevail, and the cell will elongate; for small sizes, however, it will not elongate. For $q < 0$, that is, for a consumed substance, for large sizes both diffusion forces and surface tension oppose elongation; for small sizes diffusion forces tend to elongate the cell, but the surface tension opposes it; for very small sizes the surface tension prevails. Only in a limited range of sizes can a cell elongate for $q < 0$, and that only for a special choice of constants. These results are obtained by an exact treatment of the problem of stability of a spherical cell (*MB*, chap. ix). They can be obtained in a similar way by the use of the present approximate method.

If h is very large, so that we can put in equation (14) $h = \infty$, then that equation reduces to

$$\frac{1}{r_1} \frac{dr_1}{dt} = \frac{3RT\mu}{2M\eta} \frac{\delta D_i D_e (r_1 - r_2)(\bar{c} - c_0)}{(2\delta D_i + r_1 D_e)(2\delta D_i + r_2 D_e)}. \tag{16}$$

In this case dr_1/dt is always positive for $q > 0$ and negative for $q < 0$. For simplicity we shall confine ourselves to this case in the following discussion (cf. chap. ii, p. 30).

Making $h = \infty$ in equation (18) of chapter i, introducing the expression for Λ into equation (21) of chapter i, and introducing the expression for $\bar{c} - c_0$, so obtained, into equation (16), we find

$$\frac{1}{r_1}\frac{dr_1}{dt} = \frac{RTr_1r_2q\mu}{2M\eta}\,\frac{(r_1 - r_2)\delta}{2(2\delta D_i + r_1D_e)r_1 + (2\delta D_i + r_2D_e)r_2}. \quad (17)$$

Putting $\delta = r_2$, we have

$$\frac{1}{r_1}\frac{dr_1}{dt} = \frac{RTq\mu}{2M\eta}\,\frac{(r_1 - r_2)r_1r_2^2}{2(2r_2D_i + r_1D_e)r_1 + (2D_i + D_e)r_2^2}. \quad (18)$$

Gale Young[4] worked out exactly the mathematical problem of the elongation of an ellipsoid of revolution under the influence of diffusion forces, for the case $h = \infty$, assuming that, as it elongates, it still remains exactly an ellipsoid. The equation which he obtained is rather similar to equation (18). An actual comparison shows that both equations give rather similar numerical results within a range of elongation that is practically important. For very large ratios r_1/r_2 the two equations differ widely. Although Young's formula, as we have said, is obtained by an elaborate exact solution of the problem, no preference can be attached to it because it holds exactly only for a perfect ellipsoid, a situation which is not met biologically. It is important, however, that within a practically significant range of elongations Young's exact results agree with our approximate ones, lending thus a further justification to the method used here.

Hitherto we have considered only the effect of the diffusion forces. But any cell is subject also to forces produced by surface tension. As is well known, these forces always tend to make the cell spherical, because the sphere has the smallest surface for a given volume. Thus it can be seen at once that the surface-tension forces will oppose any elongation of the

cell. The quantitative calculation of the effects of surface tension on the rate of elongation may be made approximately in the following way.

At any point on the surface the surface tension γ results in a pressure, directed inward and equal to[5]

$$-\gamma\left(\frac{1}{R_1}+\frac{1}{R_2}\right), \tag{19}$$

where R_1 and R_2 are the principal radii of curvature. For a cell like the one represented on Figure 1, we have approximately at the "ends," or "poles"

$$\frac{1}{R_1}+\frac{1}{R_2}=\frac{2}{r_2}, \tag{20}$$

while on the "sides" we have approximately

$$R_1=r_1\,; \qquad R_2=r_2\,; \qquad \frac{1}{R_1}+\frac{1}{R_2}=\frac{1}{r_1}+\frac{1}{r_2}. \tag{21}$$

We thus have at the ends a surface force equal to

$$-\frac{2\gamma}{r_2}, \tag{22}$$

and at the sides a force

$$-\gamma\left(\frac{1}{r_1}+\frac{1}{r_2}\right). \tag{23}$$

Those forces contribute to the expression I_2, given by equation (8). We now have, on the "sides,"

$$xX_\nu+yY_\nu=-\gamma r_2\left(\frac{1}{r_1}+\frac{1}{r_2}\right)$$

and, at the "ends,"

$$zZ_\nu=-\gamma r_1\frac{2}{r_2}.$$

By an argument similar to that which led to equation (13) we find the contribution of the surface tension to the quantity I_2 to be of the form $-\frac{8}{2}V\gamma(r_1 - r_2)/r_1r_2$. Therefore, according to equation (9), the contribution of the surface tension to the relative rate of elongation is

$$-\frac{\gamma}{2\eta}\frac{r_1 - r_2}{r_1r_2}. \tag{24}$$

Expression (24) can be obtained also by a different argument (*MB*, Appendix). The same expression is obtained *exactly* for the case when the shape of the cell is an ellipsoid of revolution.[4]

Adding expression (24) to the right-hand side of equation (18), we now obtain the following complete equation of elongation:

$$\left.\begin{aligned}\frac{1}{r_1}\frac{dr_1}{dt} = \frac{RTq\mu}{2M\eta}\,&\frac{(r_1 - r_2)r_1r_2^2}{2(2r_2D_i + r_1D_e)r_1 + (2D_i + D_e)r_2^2}\\ &-\frac{\gamma}{2\eta}\frac{r_1 - r_2}{r_1r_2}.\end{aligned}\right\} \tag{25}$$

By adding expression (24) to the more general equation (14) we find a corresponding more general and more complicated equation for the total relative rate of elongation. For the present we shall limit ourselves to the case of very large permeability, so that we can put $h = \infty$ in equation (14), which leads to equation (18) and through it to equation (25).

In order that elongation should take place at all, we must have $dr_1/dt > 0$. If an originally spheroidal cell should begin to elongate spontaneously, the coefficient of $r_1 - r_2$ must be positive for $r_1 = r_2 = r_0$. Under these conditions the slightest accidental deviation from the perfect spherical shape will become more and more enhanced. Putting $r_1 = r_2 = r_0$ in the right-hand side of equation (25), we find that the expression

thus obtained is negative for small values of r_0 and positive for large ones. This means that a spheroidal cell, producing a substance at a constant rate q gm cm^{-3} sec^{-1}, will not elongate if its size r_0 is less than a critical value r^*. If, because of growth, it reaches a size r^* or exceeds it, it begins to elongate spontaneously according to equation (25). The value r^* of r_0, at which that expression just changes sign, is obtained by putting it equal to zero and solving with respect to r^*. We thus obtain an expression for the critical size of a spheroidal cell, above which it elongates spontaneously. The expression thus obtained is

$$r^* = \sqrt[3]{\frac{3M(2D_i + D_e)\gamma}{RT\mu q}} \tag{26}$$

A cell does not necessarily need to increase in size in order to elongate. Its size r_0 may remain constant; but the different quantities, such as D_i, D_e, γ, and q, which determine the critical radius r^*, may vary. For a particular set of values of these quantities we may have $r_0 < r^*$, and the cell does not elongate. For another set of values r_0 may become greater than r^*, and the cell begins to elongate.

When the elongation begins, r_1 increases while r_2 decreases. If the growth of the cell is sufficiently slow, so that the volume of the cell does not vary appreciably during the process of elongation, then we may express r_2 in terms of r_1 and of the volume V from equation (10). Introducing the value of r_2 so obtained into equation (25), we obtain an equation containing r_1 only. When r_1 becomes very large and r_2 very small, so that $r_1 \gg r_2$, then, as is readily seen from equation (25), the first term of the right-hand side varies like $1/r_1$, while the second term varies like $(1/r_2) \sim \sqrt{r_1}$. Hence, denoting by A and B two constants, for large values of r_1, dr_1/dt varies like $A - Br_1^{3/2}$. But this expression is negative for sufficiently large values of r_1. This shows that for sufficiently large elonga-

tions the right-hand side of equation (25) becomes equal to zero. For the value of r_1, at which this happens, and for the corresponding value of r_2, the diffusion forces tending to elongate the cell are just equal and opposite to the surface-tension forces. The elongation therefore, once begun, proceeds only to a certain finite value. It will be shown in chapter iv that, in some cases, when for a spheroidal cell the critical size r^* is exceeded even by the smallest amount, the elongation which sets in proceeds to a finite extent, determined by the foregoing considerations. That is, for $r_1 = r_2 = r_0 > r^*$ the quantity dr_1/dt is positive and remains positive until the ratio r_1/r_2 exceeds a definite value.[6]

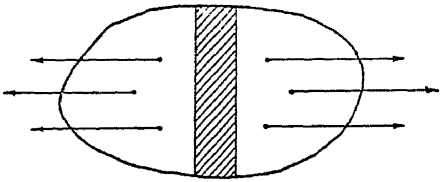


FIG. 10

Once the limiting elongation is reached, dr_1/dt becomes zero. But Betti's formula, used in derivation of equation (25), gives us the *average* rate of elongation for the cell as a whole. It may be readily seen that the elongation of the different parts of the cell will be different. In fact, at the equator the cell is pulled apart in two opposite directions by forces acting on every element of volume of each half of the cell (Fig. 10). On the other hand, any cross-section of the cell near the end is acted upon only by the force applied to each element of volume of the relatively smaller part of the cell. Thus we should expect the middle of the cell to elongate more rapidly than the ends. But, since the elongation is accompanied by a constriction in the perpendicular direction, the cell constricts more rapidly in the middle than at the ends and thus becomes

dumbbell-shaped. Even when the *average* elongation has stopped after reaching its limiting value, the cell may continue to elongate and to constrict in the middle, at the same time shortening and thickening somewhat at the ends. The actual distribution of the elongations of different parts of the cell depends on the exact shape of the latter. For the case of an ellipsoid of revolution the problem has been treated by Gale Young.[6] The theory of the process of constriction in its early stages is still wanting. For sufficiently advanced stages an approximate treatment has been given by Gale Young.[6]

When the constriction has proceeded sufficiently far, so that the cell has approximately the shape represented in

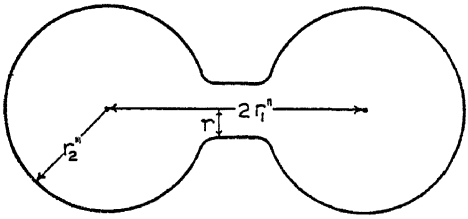


FIG. 11

Figure 11, we may estimate the further elongation of the cell by considering it as two spheres, connected by a neck and repelling each other. The repulsion is due to the fact that each sphere is producing a substance which flows outward, a force resulting which acts on each element of the other sphere in a direction away from the center of the first.

The force per unit volume at any point outside the sphere and due to that sphere is proportional to the total rate of production, that is, to qr''^3_2, and is inversely proportional to the square of the distance from the center of the sphere to that point and to the external diffusion coefficient. The force must be multiplied by the volume of the other sphere. Therefore, the whole force is directly proportional to qr''^6_2 and inversely

proportional to a^2D_e. The approximate expression for this force F, derived by Gale Young,[6] is

$$F = \frac{RT\pi\mu q r_2'^6}{6MD_e r_1'^2}. \tag{27}$$

This force may be considered as applied to the total surface πr^2 of each end of the neck, so that per square centimeter there is a force $F/\pi r^2$ directed along the line of centers outward. Besides, there is a lateral surface tension force tending to hold each end onto the neck, equal to $2\pi r\gamma$, which is again equivalent to a force

$$-\frac{2\gamma}{r} \text{ dynes cm}^{-2}, \tag{28}$$

applied to each square centimeter of the ends of the neck directed along the line of centers inward. On the lateral surface of the neck we have a capillary pressure of about

$$-\frac{\gamma}{r} \text{ dynes cm}^{-2}, \tag{29}$$

directed inward. As the spheres tend to pull apart, the narrowest region on the neck constricts inward, much like the "necking" of an iron rod in a testing-machine. Introducing the above-mentioned forces into Betti's equation and applying the latter to the "end" and "sides" of the neck alone, and ignoring the internal diffusion forces inside the neck, which in this case are rather small compared with the force F, we find, for the relative rate of change of the length l of the neck,

$$\frac{1}{l}\frac{dl}{dt} = \frac{1}{3\pi\eta}\frac{F - \pi r\gamma}{r^2}. \tag{30}$$

For a viscous incompressible body the relative lateral constriction, due to elongation, is half the relative elongation, so that we have

$$\frac{1}{r}\frac{dr}{dt} = -\frac{1}{2l}\frac{dl}{dt}. \tag{31}$$

Putting

$$P = \frac{\gamma}{6\eta}; \qquad Q = \frac{RT\mu q r_2'^6}{36\eta D_e r_1'^2 M}, \tag{32}$$

and introducing equation (31) into equation (30), we find

$$\frac{dr}{dt} = P - \frac{Q}{r} \tag{33}$$

If $Q/r \gg P$, which, with plausible values of the constants, is frequently the case, then

$$\frac{dr}{dt} = -\frac{Q}{r} \tag{34}$$

In any case, the condition $Q > Pr$ is necessary for a constriction to proceed at all. If this inequality is not satisfied, the cell will not constrict. Integrating equation (34) and denoting by r' the value of r at $t = 0$, we find

$$r^2 = r'^2 - 2Qt. \tag{35}$$

This equation shows that r becomes zero; in other words, the two half-cells separate completely at

$$t = \frac{r'^2}{2Q}. \tag{36}$$

All the formulae derived in this chapter were obtained on the assumption that a single substance is produced in the cell.

Actually, however, a very large number of substances are produced and consumed in every cell. The flow of the produced substances results in forces which tend to divide a cell. The flow of the consumed substances gives rise to forces in the opposite direction. In order to know whether the cell will divide or not, we must know the direction of the resultant of all these forces. With the very large number of metabolites involved in any cell, it is out of the question to calculate this resultant force exactly. However, some approximate estimate can be made.

From the discussions of chapter i—in particular from equation (23)—it follows that for $h = \infty$, the average gradient of concentration in a spheroidal cell is approximately proportional to qr_0/D_i. Since the average force per unit volume is $RT\mu/M$ times the gradient, it follows, therefore, that this force is proportional to q/MD_i. The same thing follows for an oblong cell from equation (24) of chapter i. If we have a large number of metabolites in the cell, those with a larger q/MD_i will contribute most to the mechanical forces. As it happens, a group of reactions in the cell usually stands out because of its particularly high rate q. This group involves the respiratory reactions discussed in chapter ii. For those reactions (consumption of sugar and oxygen, production of lactic acid and carbon dioxide) the values of q are about 10^{-6} gm cm^{-3} sec^{-1} (*MB*, chap. x). The other important reactions, such as are involved in protein breakdown, are of the order of 10^{-9} gm cm^{-3} sec^{-1}. The molecular weights of the substances involved in the respiratory reactions are relatively low, while the molecular weights of the different amino acids are higher. This makes q/M larger for the respiratory reactions. To some extent this is upset by the circumstance that we would, in general, see the heavier molecules diffuse more slowly and thus have a smaller D_i. However, this, as shall be discussed presently, is not necessarily the case. At any rate,

because of the large difference in q, it seems justifiable to discuss the case, as a first and very rough approximation, as if only the respiratory reactions were contributing to the mechanical forces. Using the same notations as in chapter ii, denoting by q_4 the rate of production of carbon dioxide, and omitting possible mechanical effects of flowing water, for reasons discussed in *MB*, chapter x, we find that the sum of the forces, owing to the four above-mentioned reactions of the respiratory group, is proportional to the quantity

$$\frac{q}{MD} = \frac{q_1}{M_1 D_{1i}} + \frac{q_2}{M_2 D_{2i}} + \frac{q_3}{M_3 D_{3i}} + \frac{q_4}{M_4 D_{4i}} . \qquad (37)$$

The quantity (37) can be expressed in terms of q_3/M_3, that is, the molar rate of the consumption of oxygen, and of the glycolytic coefficient β, introduced in chapter ii (p. 34). From the definition of β it follows that

$$\frac{q_1}{M_1} = \frac{q_3}{M_3}\left(\tfrac{1}{2}\beta + \tfrac{1}{6}\right) ; \quad \frac{q_2}{M_2} = -\frac{q_3}{M_3}\beta ; \quad \frac{q_4}{M_4} = -\frac{q_3}{M_3} . \qquad (38)$$

Introducing this into equation (37) and putting

$$\begin{aligned} \overline{A} &= -\frac{q_3}{M_3}\left(\frac{1}{D_2} - \frac{1}{2D_1}\right) ; \\ \overline{B} &= -\frac{q_3}{M_3}\left(\frac{1}{D_4} - \frac{1}{D_3} - \frac{1}{6D_1}\right) , \end{aligned} \qquad (39)$$

we find

$$\frac{q}{MD} = \overline{A}\beta + \overline{B} . \qquad (40)$$

In a previous discussion (*MB*, chap. x) we made the assumption that the diffusion coefficients are inversely proportional to the cubic roots of the molecular weights. In this case

it is found that $\overline{A} > 0$; $\overline{B} < 0$. Thus, for small values of β the expression $\overline{q/MD}$, is negative, showing that the forces due to consumed substances prevail and that the resultant force is directed inward. The assumption $D \sim M^{1/3}$ is, however, not always justified. Some indirect evidence[7] indicates that the diffusion coefficient for lactic acid is very much smaller than that for oxygen, although the size of the molecule is not so much larger. The following explanation of this may be suggested. The diffusion of different metabolites takes place in the cytoplasm, which is a complex colloidal medium. Some metabolite molecules, especially such as those of lactic acid, which have an appreciable dipole moment, may be mostly adsorbed to much larger colloidal particles. Such a molecule is then transported in the diffusion flow not as a free molecule but partly as an adsorption compound: Colloidal micella+Metabolite molecule. The diffusion coefficient in this case will be determined largely by the size of the colloidal micella and may therefore be very small. Two molecules may be of the same size, but one may be almost completely adsorbed and the other may be in a free state. Thus they may have very different diffusion coefficients. A preliminary study of such a picture has been made by John M. Reiner.[8]

Since lactic acid is more polar than glucose, it is to be expected that $D_2 < D_1$, so that $\overline{A} > 0$, since $q_3 < 0$. In that case the resulting force increases in the positive (outward) direction with increasing β. For $\overline{B} < 0$ the force is directed outward only when $\beta > -\overline{B}/\overline{A}$. For $\overline{B} > 0$ the force is always directed outward but increases in intensity with increasing β. Thus a cell should divide the more readily and the process of division should be the more rapid, the greater the glycolytic coefficient β. This, in a general way, suggests a connection with the well-known findings of O. Warburg[9] that abnormally rapidly multiplying tumor cells have a much

larger glycolytic coefficient than normally multiplying cells. While for the latter β is of the order of 0.1, for the former it may be as large as 4. Inasmuch as Warburg's findings refer also to the rate of growth, we shall discuss this important question more in detail in chapter iv.

If an equation would involve q, M, and D_i only in the combination q/MD_i, then all that we would have to do is to substitute for that combination the expression $\bar{A}\beta + \bar{B}$. However, the diffusion coefficients also enter separately in a number of formulae. Inasmuch as we are dealing here with very rough approximations, we may substitute for M, D_i, and D_e some suitably chosen average values, $\tilde{M}$ and $\tilde{D} \sim \tilde{D}_i \sim \tilde{D}_e$, and put for q an expression of the form

$$\bar{q} = A\beta + B\,. \tag{41}$$

The coefficients A and B are approximately of the order of $\bar{A}\tilde{M}\tilde{D}$ and $\bar{B}\tilde{M}\tilde{D}$, respectively. Since $\bar{A}$ and $\bar{B}$ are both proportional to q_3, as seen from equation (39), the average $\bar{q}$ in our equations will also be proportional to q_3, the rate of oxygen consumption; and in general both will be of the same order of magnitude. Except for the case where $\bar{B} < 0$ and β is just about equal to $-\bar{B}/\bar{A}$, the value of q will also be of the same order of magnitude as that of q_3. We shall call the quantity $\bar{q}$ the "effective rate of metabolism."

With these approximations we may proceed to a comparison of different equations, derived in this chapter, with experimental data. First of all, we shall discuss equation (26). The work of Newton Harvey, Kenneth Cole, and others indicates[10] that $\gamma \sim 1$ erg cm^{-2}. For room temperature $RT \sim 2.4 \times 10^{10}$ ergs. Taking, as possible values for D_e and D_i, 10^{-7} cm^2 sec^{-1} (cf. chap. ii), $\mu = 0.1$, $q \sim 10^{-6}$ gm cm^{-3} sec^{-1} and $M \sim 100$, we find the value of r^* to be of the order of 10^{-3} cm. These values are of the same order of magnitude as the average values of actual cells. It should be noted

that the relatively large variations in the values of D, q, γ, and other constants entering into equation (26) result in relatively much smaller variations of r^*. This throws some light on the understanding of the circumstance that exceptionally large or exceptionally small cells are found relatively seldom in spite of the tremendous variations in the physicochemical constitution of different cells.

Equation (26) may be compared with observations in a still different way. Since we must introduce into it for q the quantity $\bar{q}$, defined by equation (41), and since, as we have seen, $\bar{q}$ is proportional to the oxygen-consumption rate, therefore, other parameters being kept constant, the size r^* of the cell should vary as the inverse cubic root of the oxygen consumption. If we could find a number of cells, differing from each other only with respect to size and to oxygen consumption, and if we would plot the former against the latter on a logarithmic scale, we should expect that the points will fall on a straight line with the slope $-\frac{1}{3}$. It is, however, impossible to find such an ideal case. Actually, other parameters characterizing the cell, such as the diffusion coefficients, surface tension, etc., will vary from case to case. Since there is no expected correlation between these constants, we should now expect that, in plotting cell sizes against their oxygen consumption on a logarithmic scale, the points will not fall on any line at all but will *cluster* along a straight line of slope $-\frac{1}{3}$. For a proper check of this expectation we should have at least a few dozen points, so that they would form a definite cluster. Unfortunately, the data available are very meager, and only very few experimental points are available. These, however, show the proper trend, as seen from Figure 12. The legend to this figure gives the sources of the data. For cells like those of the heart tissue, which are very oblong, the average size r^* was computed from their volume V, defining r^* by $\frac{4}{3}\pi r^* = V$. This is justified by equation (26), inasmuch as we

might have plotted the volume against $\bar{q}$ and would then expect a simple inverse proportionality. More exact measurement on average cell sizes and on their oxygen-consumption

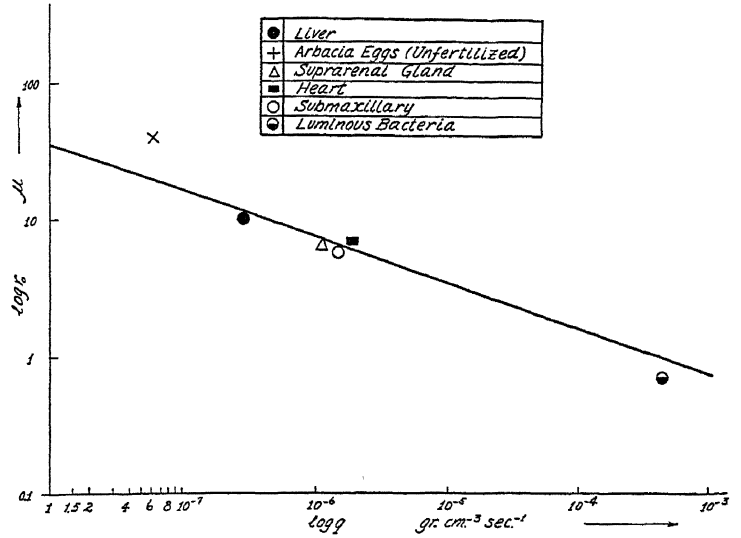


FIG. 12.—Radii r_0 of different cells plotted against q, the rate of oxygen consumption in gr cm^{-3} sec^{-1} on the logarithmic scale. The straight line has a slope $-\frac{1}{3}$. Values for q were taken from W. H. Bayliss, *Principles of General Physiology* (4th ed.), p. 612 (London: Longmans, Green & Co., 1927). Values for r_0 were taken from Stohr's *Lehrbuch der Histologie*, ed. W. von Möllendorf (22d ed.) pp. 84, 220, 283 (Jena: G. Fischer, 1930). For heart cells, an average r_0, corresponding to the average volume of the cell, is used. For values for *Arbacia* eggs and luminous bacteria, cf. Landahl.[3]

rates are urgently needed and would be of immense value for the development of the theory.

The comparison of equation (25) with experimental data presents similar difficulties because of a practically complete absence of necessary relevant observations.

K. Bělař[11] has recorded the length of the dividing nucleus

in spermatocytes of a snail for different times. A dividing nucleus undergoes a large number of structural changes; and phenomena such as movements of chromosomes, appearance of spindle fibers and others, which are not included in the present theory, take place in it. Therefore, it might appear rather doubtful as to whether there is any sense in comparing our equations with such data as Bělař's. On the other hand, however, as we have emphasized in chapter i, the approximation method, by introducing average quantities and neglecting local structural variations, leads us to relations which should be independent of those local structural details. We would therefore expect that the gross features of a dividing nucleus and of a dividing cell as a whole would be represented approximately by the same equations. Thus it is of interest to compare equation (25) with Bělař's data.

One thing, however, must be kept in mind when making such a comparison. Equation (25) is derived with the assumption that the viscous resistance and the surface tension are the only two forces opposing the diffusion forces. In a nucleus it appears highly probable that the elastic forces of the stretched spindle fibers are also an important contributing factor. These elastic forces can easily be taken into account (*MB*, p. 313) and result in the addition of an extra term to the right-hand side of equation (25). Although our knowledge of nuclear structure in mitoses is exceedingly meager, we may expect that these elastic forces will manifest themselves only in the earlier stages of the nuclear elongation, previous to the moment of complete separation of the daughter-chromosomes in the equatorial plate. According to a picture of the mechanism of mitosis suggested by the author (*MB*, chap. xiii), the elastic spindle fibers are connected at their ends to the chromosomes, on one hand, and to the body of the cell near the poles, on the other. As the nucleus elongates, these fibers are stretched as long as the chromosomes remain in the equatorial

plate. As soon as they separate, the fibers contract freely, moving the chromosomes to the poles. At this phase, however, when the fibers contract freely, the additional term in

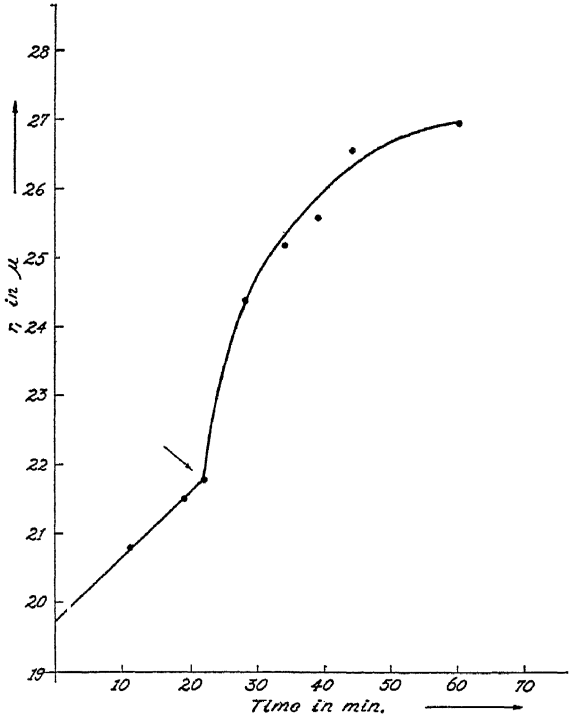


FIG. 13.—Length, r_1, of a dividing nucleus plotted against time, t. Points represent observations by K. Bělař.[11] The curve to the right of the arrow represents the theoretical equation derived by G. Young.[4] A graph of equation (44) is practically the same. At the point indicated by the arrow, the chromosomes separate. See text, p. 61.

equation (25), due to elastic forces, vanishes. Thus, on the basis of this picture or a similar one we may expect a discontinuity in the rate of elongation, just at the moment when the chromosomes separate. Up to that moment the elongation

curve should be described by an equation containing the additional elastic term. Beyond that moment the elongation should be described by equation (25).

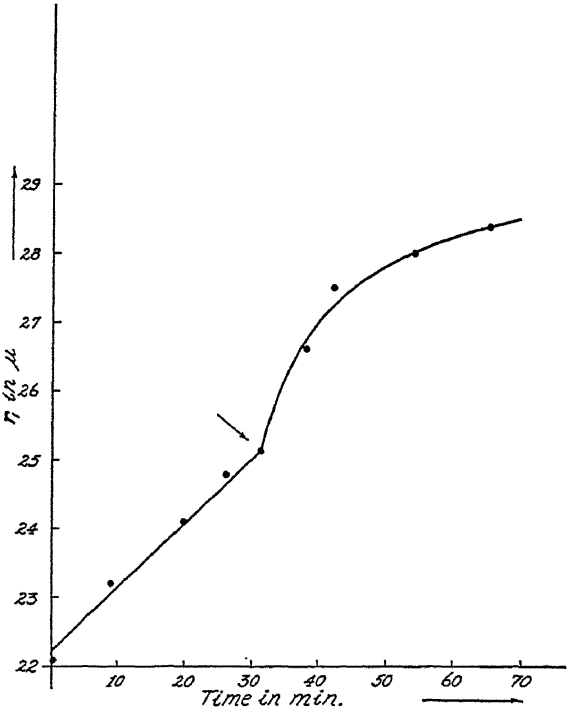


FIG. 14.—Same as Fig. 13 but for another set of data, by K. Bělař[11]

In considering the whole cell, such a discontinuity is not to be expected, owing to an absence of rigid connections of the nucleus with the cell. In fact, the dividing spindle may be moved inside the cell relatively easily.

Although these conclusions are by no means binding, it is of interest to note that Bělař's data actually show a discontinuity, as indicated by the arrow on Figures 13 and 14.

The inaccuracy of Bělař's data makes any definite conclusions impossible. We have already mentioned above (p. 46) that Gale Young[4] has derived an equation for the elongation of a cell, assuming that the latter has, during the elongation, the shape of an ellipsoid of revolution and that his equation is practically equivalent to our equation (18). By adding the effect of the surface tension, he therefore arrives at an equation equivalent to our equation (25). He has compared his equation with the experimental points lying to the right of the discontinuity in Figures 13 and 14. The graph of equation (25) is practically the same as that shown in Figures 13 and 14.

For an actual comparison of equation (25) with any data, that equation is better written in a different form. By rearranging the denominator of the first terms of the right-hand side, equation (25) can be written thus:

$$\left.\begin{aligned}\frac{1}{r_1}\frac{dr_1}{dt} = \frac{RTq\mu}{2M\eta}\,\frac{(r_1 - r_2)r_1r_2^2}{2D_e(r_1^2 + \tfrac{1}{2}r_2^2) + 4D_i(r_1r_2 + \tfrac{1}{2}r_2^2)} \\ - \frac{\gamma}{2\eta}\,\frac{r_1 - r_2}{r_1r_2}\,.\end{aligned}\right\} \quad (42)$$

Introducing the abbreviations

$$a = \frac{3V}{4\pi}\,; \quad b = \frac{2D_i}{D_e}\,; \quad K = \frac{RT\mu qa}{4M\eta D_e}\,; \quad L = \frac{\gamma}{2\eta\sqrt{a}}\,, \quad (43)$$

multiplying equation (42) by $r_1^2/(r_1 - r_2)$, and rearranging, we finally obtain

$$\frac{r_1}{r_1 - r_2}\frac{dr_1}{dt} = \frac{K}{1 + \tfrac{1}{2}(1 + b)\,\dfrac{a}{r_1^3} + b\,\dfrac{\sqrt{a}}{r_1^{3/2}}} - Lr_1^{3/2}\,. \quad (44)$$

Since a can be directly measured, the foregoing equation has three parameters: K, L, and b.

Figures 15 and 17 illustrate an application of equation (44) to the case of the division of a cell as a whole. The points on Figures 15 and 17 represent the total length of dividing isolated *Arbacia* blastomeres, taken from a micromotion film made by Professor Robert Chambers. The broken lines represent the theoretical curves, obtained by graphically integrating equation (44). For each point, r_1 was measured directly,

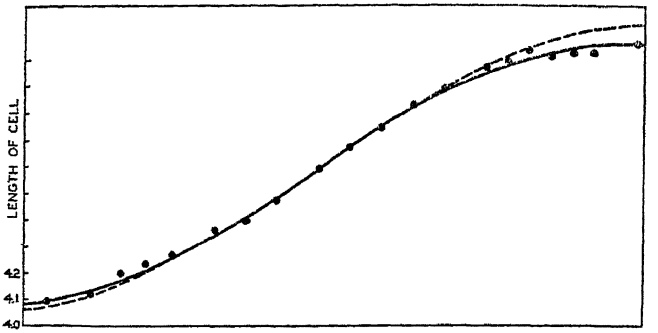


FIG. 15.—Length of a dividing isolated blastomere from the second cleavage of an *Arbacia* egg, plotted against time. The points are taken from a micromotion film by Professor Robert Chambers. The full line is the empirical curve drawn through the points. The broken line represents the theoretical equation (44). Notice that the origin of coordinates in this figure corresponds to a length of 3 (arbitrary units). The relative error of the theoretical curve is therefore rather small.

and r_2 was computed from the volume, which was determined from the initial radius of the cell, just prior to the beginning of the elongation, by using equation (10). Figures 16 and 18 represent a similar comparison of equation (35), which gives the width of the constriction as a function of time. That equation, by the method of its derivation, should hold only when the constriction has proceeded sufficiently. We should therefore expect greater discrepancies between theory and experiment for smaller times. This is actually the case, as seen from Figures 16 and 18.

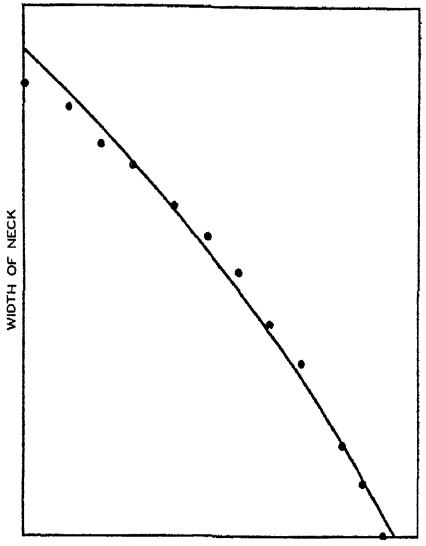


FIG. 16.—Width of constriction of a dividing isolated blastomere of an *Arbacia* egg, plotted against time. The points were taken from the same film and for the same cell as those of Fig. 15. The curve represents the theoretical equation (35).

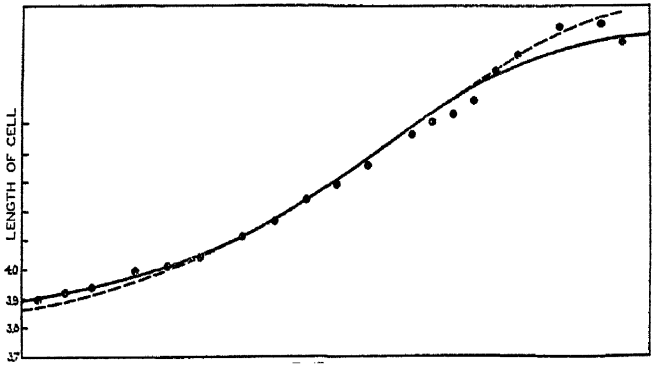


FIG. 17.—Same as Fig. 15 but for another cell, photographed on the same film.

In principle, one could attempt to calculate some of the constants of the cell from such curves as those discussed here. Unfortunately, no exact record either of the magnification or of the speeds was made by Dr. Chambers in preparing the film, since the latter was not made with the present purpose in view. Therefore, we do not attempt here to calculate the absolute values of the constants K, L, and b. As a matter of fact,

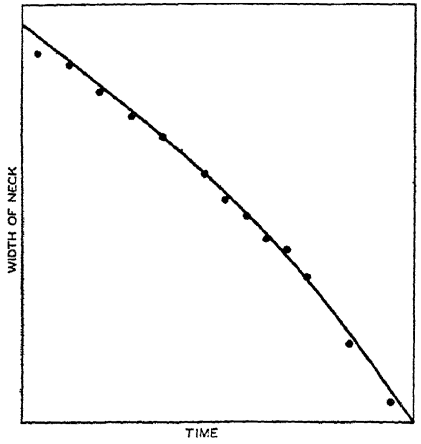


FIG. 18.—Same as Fig. 16 but for still another cell, photographed on the same film. (Same cell as used in Fig. 17.)

the whole film was not taken at regular intervals. However, the circumstance that the points, measured from every twentieth frame, do fall rather well on a smooth line, as seen on Figures 15–18, indicates that the *average* speed was relatively constant.

The following point is of interest. A look at Figure 11 shows that, approximately, $r_2'' = 0.8r_0$ and $r_1'' = 1.6r_0$. The comparison of equations (32) and (43) shows, then, that the constant Q in equation (35) is connected to the constant K in equation (44) by the relation

$$\frac{K}{Q} = \frac{88}{r_0}. \tag{45}$$

The magnification of the film used is not known exactly; but, since for a whole *Arbacia* egg the value of the radius is approximately 3.8×10^{-3} cm, the radius of the two blastomeres after the first cleavage is about 3×10^{-3} cm. Since the film used represented the second cleavage, the value of r_0 can be taken as approximately 3×10^{-3} cm. Introduced into equation (45), this gives $K/Q = 29 \times 10^3$. The actual value of K/Q for the cell corresponding to Figures 15 and 16 is found to be 24×10^3, while for the cell corresponding to Figures 17 and 18 it is 18×10^3. In view of the inaccuracy of the data, on one hand, and of the simplifications introduced in the derivation of equation (45), on the other hand, nothing more than an agreement of the orders of magnitude can be expected.

While the examples given in Figures 13–18 can be considered only as illustrations from which no definite conclusions can be drawn, they make quite obvious the importance of further exact quantitative measurements of various aspects of cell division, made under properly controlled conditions.

We have discussed in this chapter the effect of diffusion forces due to metabolism. As we see, those forces which by their nature are always present in any metabolizing cell, will produce a number of phenomena characteristic of cell division. It appears very likely that in these forces we have found the general cause of cell division, for we may say that, even if no other forces were present, cells would divide under proper conditions, because the diffusion forces are always there. However, as remarked in chapter i, a number of other forces and other factors may strongly influence and modify the actual process of cell division. Only a thorough mathematical study of all possible factors will eventually enable us to decide their relative importance. Of particular interest are, of course, possible electrical forces. N. Rashevsky[12] has pointed out that, when ions of opposite charges are metabolized in a cell,

then, because of the differences in their diffusion coefficients, the distribution of concentrations of the ions will be different and thus will give rise to the appearance of electric space charges. However, the exact treatment of this problem presents great mathematical difficulties. Using the present approximation method, Robert R. Williamson[13] has solved the problem, confirming a surmise made previously (*MB*, p. 125) that, in general, the mechanical forces, due to such electric charges, are very small, as compared to the diffusion forces discussed here.

REFERENCES

1. E. Gellhorn, *Lehrbuch der allgemeinen Physiologie*, pp. 342–61 (Leipzig: G. Thieme, 1931).

2. Gale Young, *Growth*, **2**, 165, 1938.

3. A. E. H. Love, *The Mathematical Theory of Elasticity* (2d ed.), p. 172 (Cambridge University Press, 1906).

4. Gale Young, *Bull. Math. Biophysics*, **1**, 31, 1939.

5. A. Gyemant, "Kapillarität," *Handbuch d. Physik*, **7**, 343 (Berlin: J. Springer, 1927).

6. Gale Young, *Bull. Math. Biophysics*, **1**, 75, 1939.

7. H. D. Landahl, *Bull. Math. Biophysics*, **1**, 1, 1939.

8. John M. Reiner, *Bull. Math. Biophysics*, **1**, 143, 1939.

9. O. Warburg, *Über den Stoffwechsel der Tumoren* (Berlin: J. Springer, 1926).

10. E. Newton Harvey, *Biol. Bull.*, **60**, 67, 1931, and **61**, 273, 1931; E. N. Harvey and H. Shapiro, *Jour. Cell. and Comp. Physiol.*, **5**, 255, 1934; Kenneth Cole, *Jour. Cell. and Comp. Physiol.*, **1**, 1, 1932.

11. K. Bělař, *Arch. f. Entw'mech.*, **118**, 359, 1929.

12. N. Rashevsky, *Physics*, **6**, 33, 1935.

13. Robert R. Williamson, *Bull. Math. Biophysics*, **1**, 151, 1939.

CHAPTER IV

CELLULAR GROWTH

In the previous chapter we have seen that a high glycolytic coefficient should, in general, increase the dividing forces in a cell and thus make the latter divide more readily. As has been pointed out, this agrees with the findings of O. Warburg[1] and others that cancer cells are characterized by an abnormally high glycolytic coefficient. But one of the conspicuous characteristics of tumors is their extremely rapid growth, and this is something that is not obviously connected with readier division. We shall therefore investigate in this chapter, somewhat more in detail, the manner in which the glycolytic coefficient may affect the rate of growth. And to do this we shall have to discuss the biophysical aspects of growth in general, a question which, so far, has been left outside the scope of our studies.

The increase of the mass and of the size of a cell is due to the formation of new substances, composing the body of the cell, from some simpler constituents, present in the surrounding medium and diffusing into the cell. It is therefore to be expected that the variation of the mass and size of the cell with respect to time will depend largely on the type of physicochemical reactions which we assume to be taking place within the cell. If these reactions are in some way connected with glycolysis, then the glycolytic coefficient would naturally play a part in determining the rate of growth. It might, therefore, be of interest to study such types of physicochemical reactions, and this would lead us into biochemistry rather than into biophysics. While such a study might prove to be of great importance and should be made, at present we shall

limit ourselves to a different aspect of the problem, an aspect which emphasizes more strongly the biophysical side and which shows possible *indirect* effects of glycolysis on the rate of growth of tissues. Although, unfortunately, no direct quantitative experimental data are available for comparison with the equations which we shall derive, we shall give some theoretically predicted curves which may stimulate the experimenter to carry out some measurements. Such measurements, by either confirming or refuting our expectations, will serve to decide whether effects of the type here discussed are actually occurring in some cells.

The simplest picture which we may make of the mechanism of formation of new cell material is that a number of substances penetrate by diffusion into the cell and combine there to form more complex molecules, constituting the cell body, the rate of consumption of every substance being proportional to the products of the concentrations of cell substances. To be more specific, we may, for instance, visualize each substance as an amino acid. These amino acids combine to form complex proteins. Actually, of course, the same amino acids may combine in different ways to form quite different proteins. But, for simplicity, we shall at present disregard these complications. They suggest, however, definite new types of problems, and the discussions of this chapter will perforce remain incomplete until those more complicated cases will be studied by a similar mathematical technique.

Let there be n such substances, their average concentrations inside the cell being $\bar{c}_1, \bar{c}_2, \ldots, \bar{c}_n$, and their concentrations outside the cell being $c_{01}, c_{02}, \ldots, c_{0n}$. To simplify matters, without much loss of generality, we shall consider a rounded-up, quasi-spherical cell; and we shall assume all external diffusion coefficients D_e as very large, putting $D_e = \infty$. The inner diffusion coefficient for the lth substance shall be denoted by D_{il}. For each of the substances c_l we shall have

an equation of the form of equation (21) of chapter i, in which, however, we have, since the substances are *consumed*,

$$q_l = - k_l \bar{c}_1 \bar{c}_2 \ldots . \bar{c}_n , \tag{1}$$

where k_l is a positive coefficient.

Let us put

$$k_l \bar{c}_1 \bar{c}_2 \bar{c}_3 \ldots . \bar{c}_{l-1} \bar{c}_{l+1} \ldots . \bar{c}_n = \bar{k}_l . \tag{2}$$

The left-hand side of equation (2) contains the product of all $\bar{c}$'s except $\bar{c}_l$. We now may write equation (1) thus:

$$q_l = - \bar{k}_l \bar{c}_l . \tag{3}$$

As before, we may again apply the equation for a stationary state if the rate of growth of the cell is sufficiently small, so that the cell does not alter its size appreciably during the time it takes for the diffusion state to approach stationariness. This may be done even though the diffusing substances considered here contribute directly to growth.

Equation (21) of chapter i, re-written in the notations used now, reads

$$\bar{c}_l = c_{0l} + \Lambda_l q_l . \tag{4}$$

Introducing equation (3) into equation (4), and solving with respect to $\bar{c}_l$, gives

$$\bar{c}_l = \frac{c_{0l}}{1 + \Lambda_l \bar{k}_l} , \tag{5}$$

with

$$\Lambda_l = \frac{2r_0}{9h_l} + \frac{r_0^2}{9D_{il}} . \tag{6}$$

Let us consider the case, made plausible in chapter ii (p. 30), that h is large and that therefore the first term in the

right-hand side of equation (6) may be omitted. If the molecules of the substances are rather large and strongly polar (chap. iii, p. 56), then the diffusion coefficients will be rather small—say about 10^{-9} cm^2 sec^{-1}. If the rates of reactions q_l are of the order of 10^{-6} gm cm^{-3} sec^{-1}, and the concentrations $\bar{c}_l$ are of the order of 10^{-4} gm cm^{-3}, then, as follows from equation (3), $\bar{k}_l$ must be of the order of 10^{-2} sec^{-1}. With $r_0 \sim 3 \times 10^{-3}$ cm, the expression $r_0^2\bar{k}_l/9D_{il} \sim 10 > 1$, and equation (5) reduces to

$$\bar{c}_l = \frac{9D_{il}c_{0l}}{r_0^2\bar{k}_l} \tag{7}$$

The assumptions made here are not essential and are used to simplify somewhat the final expressions.

Putting

$$\left.\begin{aligned} \sqrt[n]{D_{i1}D_{i2}\ldots.D_{in}} &= \bar{D}\,; \\ \sqrt[n]{c_{01}c_{02}\ldots.c_{0n}} &= \bar{c}_0\,; \\ \sqrt[n]{k_1\ldots.k_n} &= \bar{k}\,, \end{aligned}\right\} \tag{8}$$

we have from equations (2), (7), and (8)

$$\bar{c}_1\ldots.\bar{c}_n = \frac{9\bar{D}\bar{c}_0}{r_0^2\bar{k}}\,. \tag{9}$$

Let the substance into which the above-mentioned n substances are transformed, and which forms the body of the cell, break down at a constant rate q_b gm cm^{-3} sec^{-1}. Since the rate of formation of the body substance is equal to the sum of the rates of consumption of the n component substances, the rate of production of the body substance is equal to

$$\left.\begin{aligned} \bar{c}_1\ldots.\bar{c}_n\sum_i k_l &= k\bar{c}_1\ldots.\bar{c}_n \text{ gm cm}^{-3}\text{ sec}^{-1} \\ &\left(k = \sum_l k_l\right). \end{aligned}\right\} \tag{10}$$

Denoting by V the volume of the cell, we have for a constant density, δ, of the latter, for the rate of change of its mass δV the following differential equation:

$$\delta \frac{dV}{dt} = k\bar{c}_1 \ldots . \bar{c}_n V - q_b V . \qquad (11)$$

Introducing (9) into (11) and remembering that $V = \frac{4}{3}\pi r_0^3$, we find

$$\delta \frac{dr_0}{dt} = \frac{3\overline{D}\bar{c}_0 k}{r_0 \bar{k}} - \frac{r_0 q_b}{3} . \qquad (12)$$

Equation (12), integrated, gives,[2] denoting by C a constant of integration,

$$r_0^2 = \frac{9\overline{D}\bar{c}_0 k}{\bar{k} q_b} - Ce^{-(2q_b/3\delta)t} . \qquad (13)$$

This shows that r_0 tends asymptotically to a limiting value

$$\tilde{r}_0 = 3\sqrt{\frac{\overline{D}\bar{c}_0 k}{\bar{k} q_b}} . \qquad (14)$$

The graph of equation (13) showing the variation of r_0 with respect to time is given in Figure 19.

If, besides the reactions discussed here, respiratory and other reactions take place in a cell, with the result that above a certain size the cell will spontaneously divide in two, as discussed in the previous chapter, then two things may happen: either the critical radius r_0^*, above which division occurs, is larger than $\tilde{r}_0$, or it is smaller. In the first case the cell will reach asymptotically the size $\tilde{r}_0$ and will practically cease growing when this size is approached sufficiently closely. In the second case it will divide before it approaches $\tilde{r}_0$. As we have seen in chapter iii, the stronger the glycolytic coefficient,

the stronger the dividing forces and, everything else being equal, the smaller the r_0^*. The curve of Figure 19 shows that the smaller the cell the more rapid its rate of growth. When the cell divides at a size r_0^*, the size of the two half-cells is approximately $0.8r_0^*$ (*MB*, p. 71), and they grow up again to r_0^*. Hence, the rate of growth of the cell is always represented by that part of the curve of Figure 19 which lies between the

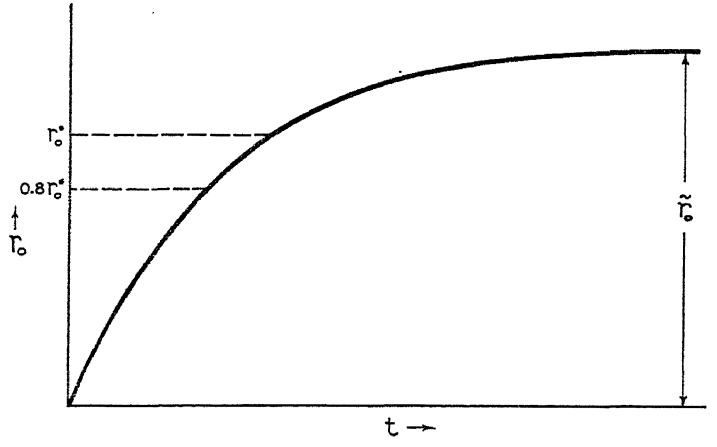


FIG. 19

ordinates r_0^* and $0.8r_0^*$. Since the slope of the curve decreases with increasing t, therefore, as is readily seen from Figure 19, the smaller the r_0^* the greater the average rate of growth between $0.8r_0^*$ and r_0^*. Therefore, the smaller the r_0^* the more frequently a cell will divide. Analytically the expression for the interval between two successive divisions is obtained by considering as the initial value of r_0^2 at $t = 0$ in equation (13) the value $0.8r_0^*$ and determining appropriately the integration constant C. Then we obtain the time at which the cell divides by making $r_0 = r_0^*$, substituting for r_0 its expression (13), and solving with respect to t.

This gives[2]

$$\Delta t = \frac{3\delta}{2q_b} \log \frac{1 - \dfrac{0.64 q_b r_0^{*2} \bar{k}}{9\bar{D}\bar{c}_0 k}}{1 - \dfrac{q_b r_1^{*2} \bar{k}}{9\bar{D}\bar{c}_0 k}} \tag{15}$$

The quantity Δt measures the time during which the number of cells in a tissue is doubled, and thus Δt can be accurately measured. Its inverse $1/\Delta t$ can properly be called the rate of growth.

For small values of r_0^*, when the second terms of the numerator and denominator are much smaller than unity, expression (15) simplifies considerably. By expanding the logarithm, by expressing r_0^* in terms of β, and by denoting the diffusion coefficients for the effective respiratory metabolism by $\tilde{D}_i$ and $\tilde{D}_e$, as in chapter iii, p. 57, we find

$$\Delta t = \frac{0.13\delta\bar{k}}{\bar{D}\bar{c}_0 k}\left(\frac{M(2\tilde{D}_i + \tilde{D}_e)}{RT\mu(A\beta + B)}\right)^{2/3}. \tag{16}$$

Equation (16) shows that, other conditions being constant, the rate of growth $1/\Delta t$ increases as the two-thirds power of the glycolytic coefficient β. According to this theory, in its simplified form, which leads to equation (16), a tissue grows the faster the greater the glycolytic coefficient β. But at the same time the average size of cells composing the tissue should be noticeably smaller. Such, however, is not generally the case for tumor tissues. This conclusion, however, does not follow from the more general equation (15), when $\bar{k}q_b r_0^{*2}/9\bar{D}\bar{c}_0 k$ is very close to unity. When it is exactly equal to 1, then Δt becomes infinite, and the rate of growth zero, because in this case $r_0^* = \tilde{r}_0$, and the critical size is reached only after an infinite time. For $q_b r_0^{*2}\bar{k}/k9\bar{D}\bar{c}_0 > 1$ the interval is imaginary.

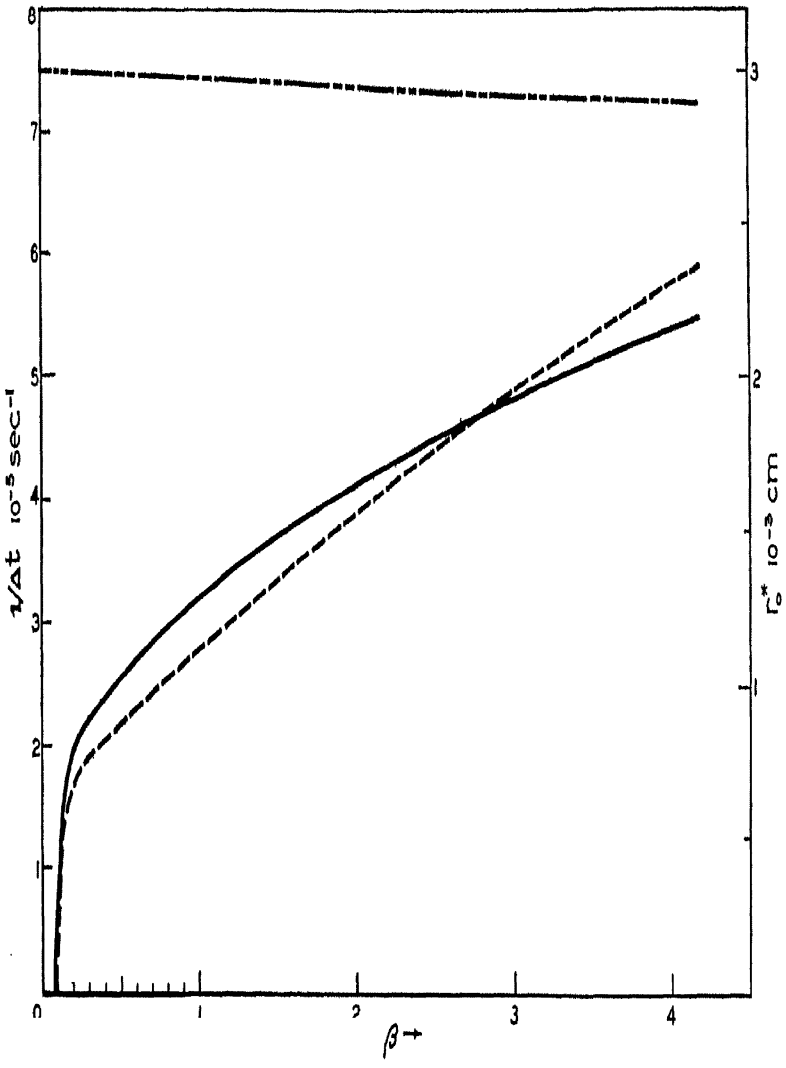


FIG. 20.—See legend on bottom of facing page

When r_0^* is still larger, so that the numerator becomes negative, then the fraction in the log of equation (15) is positive but less than 1, so that Δt is negative. But when r_0^* is slightly smaller than $\sqrt{9\bar{D}\bar{c}_0 k/\bar{k}q_b}$, then a very small variation in r_0^* will produce a very large variation of Δt. Since r_0^* varies as $q^{-1/3}$, and since q is linear in β (chap. iii), a large increase in β may result in a very small decrease of r_0^* but in a very large increase of $1/\Delta t$. In Figure 20 the full line shows the variation of $1/\Delta t$ with respect to β, calculated from equation (15), into which expression (26) of chapter iii is introduced for r_0^*. The broken line shows the variation of r_0^*.

Glycolysis may, however, influence the rate of growth in a more direct way than that considered above. When β is large, the net effective metabolic flow is directed outward and results in a pressure on the cell membrane directed outward. This pressure is given by equations (5) of chapter iii. For a

FIG. 20.—The theoretically expected variation of the rate of growth of a tissue and of the cell size with the glycolytic coefficient β. The full line represents the rate of growth, according to equation (15), with the constants: $3\delta/2q_b = 10^4$ sec; $\bar{D} = 10^{-8}$ cm^2 sec^{-1}; $c_0 = 10^{-3}$ gm cm^{-3}; $q_b = 10^{-5}$ gm cm^{-3} sec^{-1}; $\bar{k}/\bar{k} = 1$. In expressing r_0^* in terms of $\bar{q}$ and β by means of equations (26) and (41) of chapter iii, the following values were used: $A = 0.018 \times 10^{-6}$ gm cm^{-3} sec; $B = 0.738 \times 10^{-6}$ gm cm^{-3} sec^{-1}; $3\bar{M}(2\bar{D}_i + \bar{D}_e)\gamma/RT\mu = 2 \times 10^{-14}$ gm sec^{-1}. The broken line represents the variation of the rate of growth with the glycolytic coefficient β, based on equation (23). In addition to the same constants as before, the following were used: $\bar{h} = 10^{-2}$ cm sec^{-1}; $\chi = 5 \times 10^{-9}$ cm^3 dyn^{-1} sec^{-1}; $\bar{M} = 100$; $p_0 = 9.51$ dyn cm^{-2}. The alternate line represents the variation of r_0^*, which is the same in both cases. The graph does not pretend to any accuracy and is merely used as illustration. A rather high value for $\bar{h}$ has been used in order to justify the use of equations (26) of chapter iii, since the latter holds only for large values of h. Otherwise, the constants were chosen so as to give the same range of values for $1/\Delta t$ in both cases, and thus illustrate the steeper ascent of the curve corresponding to equation (23). See text, p. 80.

rounded-up cell, when $r_1 = r_2 = r_0$, the differences $c_1 - c_1'$ and $c_2 - c_2'$ both become equal to $2\bar{q}r_0/9\bar{h}$, as is seen from equations (32) of chapter i, $\bar{q}$ being defined by equation (41) of chapter iii. Thus, the pressure in this case is equal to

$$\frac{2}{9}\frac{RT}{\overline{M}}\frac{\bar{q}r_0}{\bar{h}}, \tag{17}$$

where $\bar{h}$ again refers to an average value for respiratory metabolites. The pressure, represented by expression (17), may be partially compensated by a static osmotic pressure p_0, owing to different nonmetabolized substances dissolved in the cell and in the surrounding medium—substances for which the membrane may be impermeable. The total pressure is then equal to

$$\frac{2}{9}\frac{RT}{\overline{M}}\frac{\bar{q}r_0}{\bar{h}} - p_0. \tag{18}$$

Let a reversible reaction take place in the cell, such as

$$A_1 + A_2 + \ldots . A_n \leftrightarrows P_1 + P_2 + \ldots . P_m, \tag{19}$$

where A_i represents some amino acids and P_i represents proteins. If this reaction requires a special enzyme, which is present only inside the cell, then it will not take place outside the cell. If the membrane is permeable to A_i but not to P_i, then the concentrations of A_i will be the same inside the cell and outside; and if the cell size remains constant, the amounts of P_i will not change, everything remaining static. However, let the cell size increase *very slowly* as a result of some causes. Then the concentrations of A_i and P_i will decrease. But, because of the supply from outside, the concentration of A_i will be practically immediately re-established. Thus the concentration of A_i will remain the same and that of

P_i will decrease. As a result of this, some P_i will be formed according to equation (19). The cause of the increase may be the pressure given by expression (18), which tends to expand the cell. According to this picture, a cell does not grow in size because of increase in mass, but rather increases in mass because of growth in size.

In the first approximation the rate of increase in size will be proportional to the expression (18) and will be given by

$$\frac{dr_0}{dt} = \chi\left(\frac{2}{9}\frac{RT}{\overline{M}}\frac{\bar{q}r_0}{\bar{h}} - p_0\right), \tag{20}$$

where χ is principally determined by the resistance of the membrane to the flow of water, which, because of its incompressibility, must flow into the cell whenever the latter expands. Equation (20), integrated, gives

$$r_0 = \frac{9}{2}\frac{\overline{M}\bar{h}p_0}{RT\bar{q}} + Ce^{(2RT\chi\bar{q}/9\overline{M}\bar{h})t}. \tag{21}$$

For β just above the critical value $\beta_0 = -B/A$, when $\bar{q} = A\beta + B$ is positive and very small, such a cell might grow to tremendous dimensions, though the rate of growth would be extremely slow. For normal β it will grow to a normal size r_0^* and then divide. Again, as before, let us put $t = 0$, when $r_0 = 0.8r_0^*$. We then obtain[2]

$$r_0 = \frac{9\overline{M}\bar{h}p_0}{2RT\bar{q}} + \left(0.8r_0^* - \frac{9\overline{M}\bar{h}p_0}{2RT\bar{q}}\right)e^{(2RT\chi\bar{q}/9\overline{M}\bar{h})t}. \tag{22}$$

The interval Δt between two successive divisions is determined from

$$\frac{9\overline{M}\bar{h}p_0}{2RT\bar{q}} + \left(0.8r_0^* - \frac{9\overline{M}\bar{h}p_0}{2RT\bar{q}}\right)e^{(2RT\chi\bar{q}/9\overline{M}\bar{h})\Delta t} = r_0^*,$$

or, solving it for Δt,

$$\Delta t = \frac{9\overline{M}\bar{h}}{2RT\chi\bar{q}} \log \frac{9\overline{M}\bar{h}p_0 - 2RT\bar{q}r_0^*}{9\overline{M}\bar{h}p_0 - 1.6RT\bar{q}r_0^*}. \tag{23}$$

For very small values of p_0, equation (23) reduces to

$$\Delta t = \frac{9\overline{M}\bar{h}}{2RT\chi\bar{q}} \log 1.25 = \frac{\overline{M}\bar{h}}{RT\chi\bar{q}}. \tag{24}$$

Again expressing the critical size r_0^* in terms of the glycolytic coefficient β, by means of equations (26) and (41) of chapter iii, we obtain from equation (24) the interval Δt in terms of β. Such a relation, obtained from equation (24), is shown in Figure 20 by the broken line.

Unfortunately, no quantitative data are available for comparison with the calculated curves. We see from Figure 20 that a variation of β from 0.1 to 4 results in a tremendous increase of the rate of growth. It is known empirically[1] that values of β about 1 correspond to benign tumors, while values such as 3 or 4 correspond to definitely malignant, very much more rapidly growing tumors. The importance of exact quantitative experimental studies of the rate of growth of different tumors, together with simultaneous determinations of the glycolytic coefficients of *the same tumors*, becomes apparent from these considerations. By comparing the results of such experiments with the two curves of Figure 20, either we may be able to decide in favor of the one or of the other of the equations here derived, or else we may get definite indications for the necessity of considering more complicated mechanisms involving the effects of glycolysis on the rate of growth.[2]

A careful, well-controlled quantitative statistical study of cell sizes in tumors, as compared to normal tissues, is also

urgently needed. From Figure 20 we see that the change in cell size of only a small percentage may correspond to a tremendous increase in the rate of growth. It is usually accepted that tumor cells do not show any appreciable differences in size from normal cells. But this is based on rather crude general observations. A difference of a small percentage will escape notice unless a thorough statistical study of size distributions is made and mean values and standard deviations computed and compared.

REFERENCES

1. O. Warburg, *Über den Stoffwechsel der Tumoren* (Berlin: J. Springer, 1926).

2. N. Rashevsky, *Bull. Math. Biophysics*, 1, 119, 1939.

CHAPTER V

CELLULAR FORMS AND MOVEMENTS

In chapter iii we have studied the deformation of the cell under the influence of diffusion forces for the case where the rate of deformation is very slow. In this case we may consider the diffusion field at any moment as quasi-stationary. Since in some cases the cell division lasts only 10 minutes or less, it is important to discard the restriction of slow rate of deformation and to attack the problem in its more general aspects.

In this case the variation of the concentration $\bar{c}$ in the cell is given by equation (19) of chapter i. However, the quantity Λ, which is a function of the cell shape, now itself is variable, its variation being determined by the diffusion forces according to the discussion of chapter iii. These diffusion forces are themselves determined by the values of $\bar{c}$, c_1, c_2, etc.

For the sake of simplicity, we shall confine ourselves to the case of a very large permeability, h, though the more general case is treated in the same manner. We also consider, again for sake of definiteness, the production of a substance; that is, we take $q > 0$.

Making $h = \infty$ and $\delta = r_2$ in equation (18) of chapter i, introducing the expression for Λ into equation (19), and introducing a new variable

$$c = \bar{c} - c_0 , \tag{1}$$

we may write equation (19) of chapter i in the form

$$\frac{dc}{dt} = q - \frac{3D_iD_e}{r_1r_2^2} \frac{r_2^2(D_e + 2D_i) + 2(r_1D_e + 2r_2D_i)r_1}{(r_1D_e + 2r_2D_i)(D_e + 2D_i)} c . \tag{2}$$

The rate of elongation of the cell is given by equation (16) of chapter iii, which in that form holds for the general non-stationary case. Putting in that equation $\delta = r_2$ and adding to the right-hand side the surface-tension term (chap. iii, Eq. [24]), we obtain

$$\left.\begin{aligned}\frac{1}{r_1}\frac{dr_1}{dt} = \frac{3RT\mu}{2M\eta} \\ \times \frac{D_i D_e (r_1 - r_2)}{(2r_2 D_i + r_1 D_e)(2D_i + D_e)}\, c - \frac{\gamma}{2\eta}\frac{r_1 - r_2}{r_1 r_2}\end{aligned}\right\} \quad (3)$$

We shall now discuss the simultaneous differential equations (2) and (3). These equations are nonlinear, and their exact solution presents very great difficulties. We shall, however, study the general properties of their solutions by means of a semigraphical method which is rather simple and useful.

The requirement

$$\frac{dc}{dt} \gtreqless 0\,, \quad (4)$$

introduced into (2), gives

$$c \gtreqless \frac{r_1 r_2^2 q}{3D_i D_e}\,\frac{(r_1 D_e + 2r_2 D_i)(D_e + 2D_i)}{2r_1(r_1 D_e + 2r_2 D_i) + r_2^2(D_e + 2D_i)}\,. \quad (5)$$

The equality sign in (5) gives the equation of a curve for which $dc/dt = 0$. Since

$$r_2^2 = \frac{3V}{4\pi r_e}\,; \qquad V = \tfrac{4}{3}\pi r_0^3\,; \qquad r_0 = \left(\frac{3V}{4\pi}\right)^{1/3}, \quad (6)$$

the equation (5) may be written

$$\left.\begin{aligned}c = \frac{Vq(D_e + 2D_i)}{4\pi D_i D_e} \\ \times \frac{D_e r_1 + 2D_i\sqrt{\dfrac{3V}{4\pi r_1}}}{2D_e r_1^2 + 4D_i\sqrt{\dfrac{3Vr_1}{4\pi}} + \dfrac{3V(D_e + 2D_i)}{4\pi r_1}}\,.\end{aligned}\right\} \quad (7)$$

For very large values of r_1 we have

$$c = \frac{Vq(D_e + 2D_i)}{8\pi D_i D_e} \frac{1}{r_1}. \tag{8}$$

For very small values of r_1 we have

$$c = \frac{2q}{3D_e} \sqrt{\frac{3Vr_1}{4\pi}}. \tag{9}$$

From (8) and (9) we see that $c = 0$ for $r_1 = 0$ and for $r_1 = \infty$. The value of c is everywhere positive. From physical con-

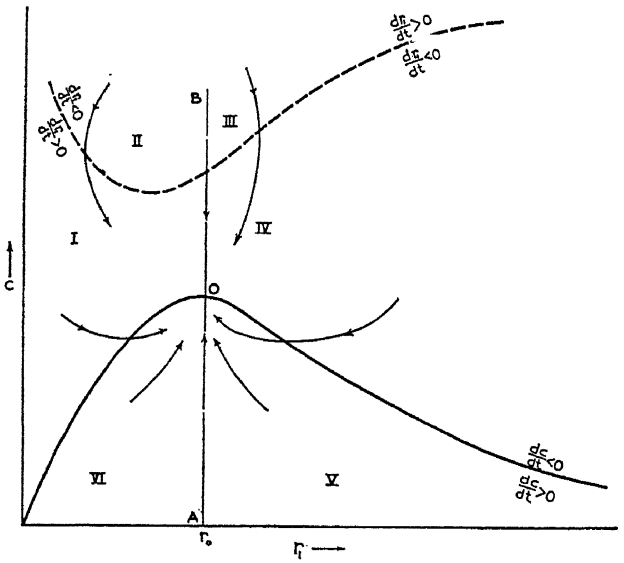


FIG. 21

siderations it is clear that c should have a maximum for $r_1 = r_0 = \sqrt[3]{3V/4\pi}$, that is, when the cell is spherical. For $D_e = \infty$ this is actually the case, as can easily be seen.

For finite D_e (7) gives a maximum at a point slightly different from r_0. This is due to the approximation introduced in chapter iii by making $\delta = r_2$, which is not quite exact.

Let us consider r_1 and c as rectangular coordinates in a plane (Fig. 21). Equation (7) then represents a line, with a maximum at approximately $r_1 = r_0$ (Fig. 21, full line). For all points of the plane which lie below that line, $dc/dt > 0$. In other words, if we choose any pair of values r_1 and c, that is, a given elongation and a given concentration, such that the representative point is below the line, then for such a pair of values the concentration $c = \bar{c} - c_0$ will increase. The reverse holds for all points above the line. Here we have $dc/dt < 0$.

Because of equation (3) the requirement

$$\frac{dr_1}{dt} \gtreqless 0 \tag{10}$$

leads to

$$c \gtreqless \frac{M\gamma(2D_i + D_e)}{3RT\mu D_i D_e}\left(\frac{2D_i}{r_1} + \frac{D_e}{r_2}\right) \tag{11}$$

for

$$r_1 > r_2 ,$$

and to

$$c \lesseqgtr \frac{M\gamma(2D_i + D_e)}{3RT\mu D_i D_e}\left(\frac{2D_i}{r_1} + \frac{D_e}{r_2}\right) \tag{12}$$

for

$$r_1 < r_2 .$$

Because of (6), equations (11) and (12) may be written thus:

$$c \geqq \frac{M\gamma(2D_i + D_e)}{3RT\mu D_i D_e}\left(\frac{2D_i}{r_1} + \frac{D_e}{r_0^{3/2}}\sqrt{r_1}\right) \quad \text{for} \quad r_1 > r_2 \tag{13}$$

and

$$c \leqq \frac{M\gamma(2D_i + D_e)}{3RT\mu D_i D_e} \quad \frac{2D_i}{r_1} + \frac{D_e}{r_0^{3/2}}\sqrt{r_1} \qquad \text{for} \qquad r_1 < r_2. \quad (14)$$

For $r_1 = r_2$ equation (3) gives always $dr_1/dt = 0$. The equality sign in (13) and (14) gives us c as a function of r_1. This function is infinite for $r_1 = 0$ and varies for small values of r_1 like $1/r_1$; while for larger values of r_1 it varies as $\sqrt{r_1}$ (Fig. 21, broken line). It has a minimum for

$$r_1 = \left(\frac{4D_i}{D_e}\right)^{2/3} r_0 . \quad (15)$$

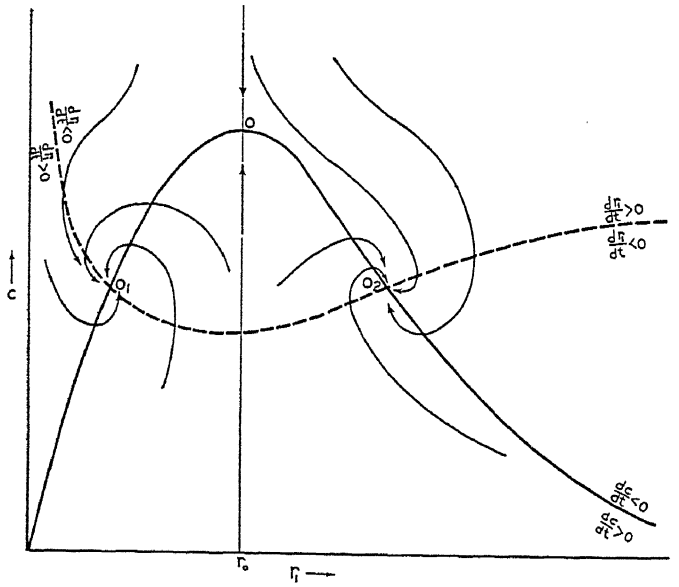


FIG. 22

The two lines (7) and (13) either may not intersect at all (Fig. 21) or may intersect as in Figure 22 or in Figure 23.

The situation shown in Figure 21 will occur when, for a given r_0 and other constants, the value of q is sufficiently small, resulting in too low a concentration $c = \bar{c} - c_0$ at the maximum, which corresponds to a sphere. In this case the two lines represented by equations (7) and (13), together with the line AB, which corresponds to $r_2 = r_0$, divide the quad-

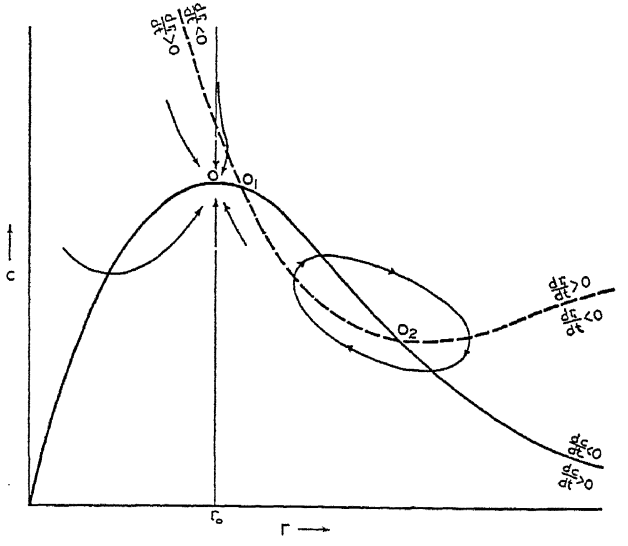


FIG. 23

rant of the r_1c plane into six regions. In region I, $dc/dt < 0$, while $dr_1/dt > 0$. If we consider a cell, characterized by a given elongation r_1 and a given concentration $c = \bar{c} - c_0$, and if to these particular values of r_1 and c the corresponding point lies in region I, then r_1 will increase while c will decrease. We may say that at a certain moment any configuration point in the r_1c plane moves in the direction indicated by the arrows. In region II we have $dc/dt < 0$ and $dr_1/dt < 0$; the direction of the movement of a configuration point is also

indicated by the arrow. In region III, $dr_1/dt > 0$ and $dc/dt < 0$; while in region IV, $dr_1/dt < 0$, $dc/dt < 0$. A similar survey of the other regions leads us to a distribution of pathways along which the configuration point moves according to equations (2) and (3), as shown in Figure 21. For any point on AB we have $dr_1/dt = 0$. Therefore, above the heavy line any such point will move downward; below it, upward. All paths converge to the point O, which corresponds to $r_1 = r_2 = r_0$, that is, to a sphere. This is quite clear physically, since for very small values of q the sphere is a stable shape of equilibrium.

A similar analysis of the case represented in Figure 22 leads to a different distribution of the paths. The point O is now a point of unstable equilibrium. There are now two points, O_1 and O_2, of stable equilibrium, toward which the paths converge. The point O_1 corresponds to $r_1 < r_0$, or $r_1 < r_2$—in other words, to a flattened shape. The point O_2 corresponds to $r_1 > r_0$ or $r_1 > r_2$—that is to say, to an elongated shape. This happens when, owing to a sufficiently large q, the sphere becomes unstable for a given r_0, that is, when r_0 becomes greater than the critical radius r^* given by equation (26) of chapter iii. In this case the metabolic forces begin to elongate the cell. Either this elongation may proceed in one direction only, leading to oblong shapes with an eventual division in two; or it may proceed in two transverse directions, leading to flattened shapes, like, for instance, red blood corpuscles.

In the case represented in Figure 23 the spheroidal shape, corresponding to $r_1 = r_0$, is in a relatively stable equilibrium. It requires a finite disturbance to make the cell elongate to the point O_2.

Around the point O_2 the path of the configuration point may be a closed line. This means that the cell may periodically contract and expand around a configuration of equilibrium.[1] Let us now keep all other parameters in equation

(7) constant but increase V or, what amounts to the same, increase $r_0 = (3V/4\pi)^{1/3}$. The point O in Figure 23, corresponding to the maximum of the heavy line, will shift to the right and will also move upward, since with increasing V the maximum concentration c increases. Finally, O will fall on the broken line, so that O and O_1 will coincide. In other words, the sphere will become a configuration of unstable equilibrium, and any infinitesimal elongation will bring the cell into configuration O_2. The value of r_0 for which this happens is nothing but the value r_0^* given by equation (26) of chapter iii. The foregoing reasoning proves the assertion made in chapter iii, page 50, that, in some cases when the spheroid shape becomes unstable, the deformation proceeds at once to a finite value.

A similar situation will occur for the case in which the maximum of the heavy line is to the right of the minimum of the broken line. In this case, however, as soon as instability for spheroidal shapes will set in, the cell will flatten out by a finite amount and assume the configuration corresponding to the point O_1 of Figure 22.

For very slow deformations, such as studied in chapter iii, the configuration point of the cell in the (r_1, c) plane moves along the heavy line, which gives at each point the value of $c = \bar{c} - c_0$, corresponding to a quasi-stationary value for a given elongation, r_1. In general the configuration point moves along one of the pathways indicated by the arrows.

Considering, for instance, the case where $D_i \gg D_e$, so that D_e may be neglected, as compared with D_i, when occurring together as a sum or a difference, and considering, moreover, the case where the point O_2 is very much to the right of O, we may use,[1] instead of equation (7) the approximate equation (8). The latter now simplifies to

$$c = \frac{Vq}{4\pi D_e} \frac{1}{r_1}, \tag{16}$$

and equation (13) now simplifies to

$$c = \frac{2M\gamma}{3RT\mu D_e}\left(\frac{2D_i}{r_1} + \frac{D_e}{r_0^{3/2}}\sqrt{r_1}\right). \tag{17}$$

By equating the right-hand sides of equations (16) and (17) we find the abscissa of the point O_2. Making use also of equation (26) of chapter iii, and putting

$$V^* = \tfrac{4}{3}\pi r_0^{*3}; \qquad \alpha = 3\,\frac{V}{V^*} - 2\,, \tag{18}$$

we find approximately[1]

$$r_1'' = r_0\left(\frac{aD_i}{D_e}\right)^{2/3}. \tag{19}$$

In a similar way we find the value c_2 of c corresponding to the point O_2.

If the condition discussed in chapter iii, page 53, is satisfied, then the cell not only will elongate but also will divide. Otherwise it remains in an elongated nonspherical shape indefinitely.

Since by means of equation (26) of chapter iii, r_0^* is expressed in terms of the effective metabolic rate $\bar{q}$, equation (19) gives us the elongation of a cell in terms of the effective metabolic rate.

We have pointed out before (*MB*, chap. ix) that, because of the presence of metabolic forces, we should expect the possibility of existence of nonspherical shapes of equilibrium in liquid cells. We also pointed out that, inasmuch as the existence of such shapes is tied up to the presence of the metabolic forces, we should expect that a cessation of metabolism—in other words, the death of the cell—would result in a rounding-up of the cells unless there are some rigid structures

present within the cell. This agrees with the general observations that a number of cells which have an oblong shape dur-

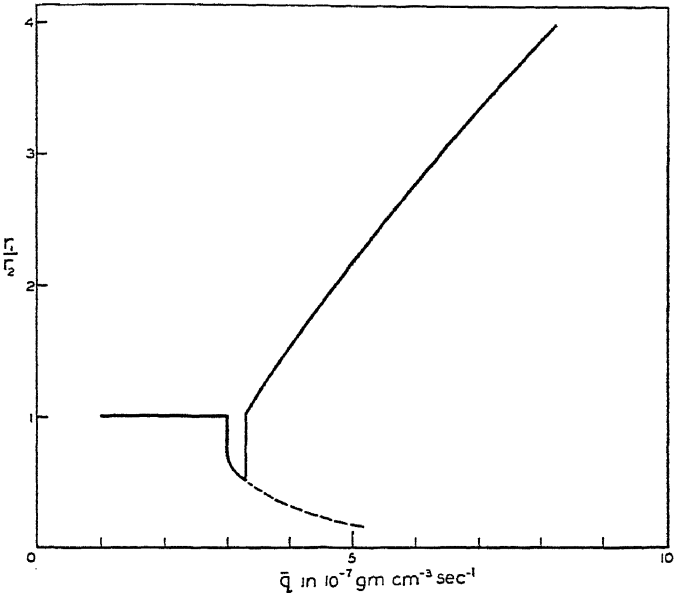


FIG. 24.—The theoretically expected variation of the length-to-width ratio r_1/r_2 of a unicellular organism with varying rate of metabolism $\bar{q}$, when the latter is influenced by external agencies, for the following values of constants: $D_e = 10^{-6}$ cm^2 sec^{-1}; $D_i = 10^{-7}$ cm^2 sec^{-1}; $\gamma = 1$ erg cm^{-2}; $\mu = 0.1$; $2RT/M = 5 \times 10^9$ erg gm^{-1}. The case corresponds to that of Fig. 22. The discontinuity is due to the fact that, for that particular value of $\bar{q}$, the configuration O_2 on Fig. 22 becomes unstable and the cell assumes the flattened shape, corresponding to O_1 of Fig. 22. If $\bar{q}$ is again increased, the cell flattens out further (broken line). We have, thus, a hysteresis effect. Such an effect is not present for the choice of constants, which correspond to Fig. 23.

ing life round up upon death, as discussed in *MB*, chapter ix. But, as has been pointed out, we should expect not only a complete rounding-up of the cell upon death but also a

gradual rounding-up during life if the cellular metabolism is impaired by some drugs (*MB*, p. 103). Although this seems to be the case, no definite quantitative observations are available, however. The discussion of this chapter leads us to a definite prediction concerning the dependence of the elongation r_1'' of a cell as a function of the effective metabolic rate $\bar{q}$. The use of equation (19) in this case is not justified, for that equation holds only for a rather narrow range of values of r_1''. We can, however, easily construct r_1'' as a function of q graphically, by plotting the lines given by equations (7) and (13) for different values of q, by using a given set of values for the other constants, and by determining graphically the abscissa of the point O_2 for every value of q. The result of such a determination is shown in Figure 24.

Since the effective q is itself a function of the other components of metabolism, we may thus study the effects of those components on the shape of the cell. The urgent necessity of studying the effects of various drugs and other environmental conditions affecting metabolism on the shape of protozoa becomes evident.

As we have said, around the point O_2 the cell may exhibit periodical contractions and expansions. This case has been studied approximately analytically,[1] and the bearing of it on a possible biophysical theory of such phenomena as contractility of tissues and amoeboid movements has been discussed.

REFERENCES

1. N. Rashevsky, *Bull. Math. Biophysics*, 1, 47, 1939.

CHAPTER VI

PROTOPLASMIC STREAMING

In all cases hitherto discussed the diffusion forces were directed either outward or inward, tending either to disrupt or to compress the cell. It is conceivable that in some cases a sufficiently asymmetrical distribution of these forces inside the cell will produce not merely elongations or contractions of the cell but continuous displacements of different parts of the cell with respect to each other, and thus produce internal movements in the cell. Such internal movements are, as a matter of fact, frequently observed in actual cells. They form a wide class of phenomena, known as "protoplasmic streaming."[1] While in plant cells these protoplasmic streamings occur rather continuously during the normal life of the cell, in some animal cells they appear in different forms during the process of cell division.

As mentioned above, some sufficiently asymmetric distribution of the forces must be present in order to induce such movements. Such an asymmetric distribution is, however, quite likely to occur whenever the cell exhibits gross inhomogeneities of its structure, which is more a rule than an exception. Thus, for instance, in most cells the nucleus does not occupy a central position but lies eccentrically. Quite regardless of any detailed assumptions we make about the chemical and physiological role of the nucleus, if we consider that it is a seat of some important reactions, we see that an asymmetry of location of the nucleus will result in an asymmetry of the reaction rates at various parts of the cell. In other words, instead of having an average constant rate of reaction q per unit volume, we now have the situation in which the rate

of reaction for *some* metabolites will be much larger near one end of the cell than near the other. If this is the case, then the concentrations of the corresponding metabolite will not be symmetric with respect to the equator of the cell. Considering again, for sake of definiteness, the production of one substance, we can readily see that, if in the region AB (Fig. 25) the rate of production is higher than in the region $A'B'$, then the concentration near the end AB will also be greater than the concentration near the end $A'B'$. Instead of having

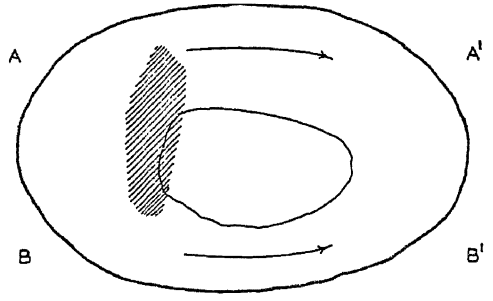


FIG. 25

a maximum value near the center of the cell and dropping down toward *both* ends, the concentration now will have a maximum somewhere near that indicated in Figure 25 by the shaded area and will drop toward the end $A'B'$. A similar thing will happen if the rate of reaction q is everywhere the same but the permeability of the cell membrane at AB is much less than the permeability at $A'B'$. Such a difference in permeability may also be caused by an asymmetric location of the nucleus if the latter produces substances which affect the permeability. There will therefore be a force directed from left to right, as indicated by arrows. The force represented by the upper arrow tends to set the interior of the cell into clockwise motion; the force represented by the lower arrow

tends to set it into counterclockwise motion. In order that the interior of the cell may be maintained in a steady motion, either clockwise or counterclockwise, continuous work must be done by the driving forces, because, owing to the viscosity of the cytoplasm, any continuous motion is connected with a dissipation of energy. This implies that the work done by the forces in carrying a particle of the cell interior around a closed path would be finite. As we have seen (chap. iii), the diffusion forces may be considered as being derived from a potential which is proportional to the concentration. In that case, however, the integral of the force along any closed circuit is always zero. The circumstance that now we may have a diffusion in a moving medium does not alter this situation. The motion of the medium will affect the distribution of the concentrations, making the gradients different, but the force still remains derived from a potential. Thus, at first sight it would appear that the diffusion forces cannot produce any permanent circulatory motion of the cytoplasm. A closer examination, however,[2] shows that this is not necessarily the case. The foregoing argument holds if we consider that the whole of the diffusion force is sustained by the liquid medium. If, however, there are some immobile rigid structures present, then part of the force will be sustained by those, and that part will not contribute to the motion of the liquid. Either these solid structures may be constituted by a jelly-like meshwork, through which the liquid may be "squeezed through," or the rigid walls of the cell may sustain part of the force, which will be transmitted to them by the viscosity of the liquid.

Suppose, now, that the physical constitution of the rigid framework or of the walls is different in the top and in the bottom part of Figure 25, so that in the two regions different fractions of the total force are sustained by the liquid. Then, while the two *total* forces, represented by the arrows, will be

equal, their parts sustained by the liquid will not be equal. The total integral of *that part* of the force along a closed circuit will not be zero, and streaming will result.

Gale Young has developed a theory of protoplasmic streaming along those lines.[2] Referring to Young's original paper for a more exact treatment, we shall give here merely an approximate outline of the mathematical argument, which leads to somewhat similar results.

We may estimate the order of magnitude of the upper limit of the velocity of streaming in the following manner. Suppose, for definiteness, that the asymmetry of the cell, represented schematically by Figure 25, is such that a greater portion of the force represented by the upper arrow is transmitted to the liquid, so that that force prevails and the liquid is set in clockwise circulation. The strongest driving force would be obtained if the force represented by the upper arrow would be completely transmitted to the liquid, while the force represented by the lower arrow would be completely sustained by the solid structure. In that case the total driving force would be represented by the upper arrow. According to equation (3) of chapter iii, this force is of the order of magnitude of $(RT/M) \times$ grad c per unit volume. Denoting the concentrations at AB and at $A'B'$ by c_1 and c_2, respectively, we find that the average value of grad c is of the order of $(c_1 - c_2)/2r_1$. Hence, the driving force per unit volume is of the order of

$$\frac{RT(c_1 - c_2)}{2Mr_1} \tag{1}$$

Considering, for simplicity, the case where the asymmetry of concentration at the two ends is due to a very low permeability at AB and a very high one at $A'B'$, we can put $c_2 = c_0$, where c_0 is the external concentration.

We shall now express $c_1 - c_2 = c_1 - c_0$ in terms of the re-

action rate q. Besides the transport by diffusion, a transport of substance by convection takes place. If v denotes the average velocity of streaming, and if $\bar{c}$ is the average concentration in the cell, then the convection transport is equal to $v\bar{c}$ gr cm^{-2} sec^{-1} (*MB*, p. 13). Remembering that the velocity has a different direction in the upper and in the lower part of the cell, and that the cross-section of each part is approximately $(\pi/2)r_2^2$, we find that the net transport from left to right is

$$\left.\begin{aligned}\left(D_i\frac{c_1 - c_0}{2r_1} + v\bar{c} + D_i\frac{c_1 - c_0}{2r_1} - v\bar{c}\right)\frac{\pi}{2}r_2^2 \\ = D_i\frac{c_1 - c_0}{2r_1}\pi r_2^2 .\end{aligned}\right\} \quad (2)$$

This net transport is of the order of magnitude of the total production rate $\frac{4}{3}\pi r_1 r_2^2 q$. Strictly speaking, we must also take into account the loss of substance through the "sides" AA' and BB'. To neglect this introduces a rather large error but does not affect the order of magnitude, which is all that we are now concerned with. Equating expression (2) to the total rate of production and solving for $c_1 - c_0$, we find

$$c_1 - c_0 = \frac{8}{3}\frac{qr_1^2}{D_i} . \quad (3)$$

Combining this with expression (1), we find the driving force per unit volume to be of the order of

$$\frac{4RT}{3MD_i}qr_1 . \quad (4)$$

The total force is equal to expression (4) times approximately half the volume of the cell, that is, times $\frac{2}{3}\pi r_1 r_2^2$. When there is a steady streaming, then the total force must be equal to

the viscous resistance. Denoting by η the viscosity, and remembering that the average "gradient of velocity" is $2v/r_2$ (because over the distance $2r_2$ the velocity changes from $+v$ to $-v$), we find the resistance per unit area to be $2v\eta/r_2$. The total resistance is obtained by multiplying this by the total surface, which is of the order of $4r_1r_2$. Now, by equating the total resistance to the total force, we obtain a linear equation for the determination of v. Solving it, we find

$$v \sim \frac{\pi RTqr_1r_2^2}{9MD_i\eta}. \tag{5}$$

Taking $q \sim 10^{-6}$ gm cm^{-3} sec^{-1}, $r_1 \sim r_2 \sim 10^{-3}$ cm, $D_i \sim 10^{-7}$ cm^2 sec^{-1}, and $\eta \sim 1$, that is, a value a hundred times that of water, we find $v \sim 1$ cm sec^{-1}. The actual velocities[3] are of the order of $10^{-3} - 10^{-5}$ cm sec^{-1}. This is not surprising, for we have given here the upper limit of the velocity, under the assumption that the difference of the two oppositely acting forces (Fig. 25) is simply equal to one of them, the other force being zero. The fact that the actual velocities are so much smaller than the upper limit shows that a very small asymmetry, resulting in a very small difference in the amounts of the forces sustained by the solid structure, will be sufficient to produce streaming.

Another interesting consequence of equation (5) is the following: The velocity v depends both explicitly and implicitly on the temperature. The temperature-depending quantities on the right-hand side of equation (5) are q, D_i, and η. Inasmuch as D_i is approximately[4] inversely proportional to η, being, besides, proportional[4] to T, the product $D_i\eta$ is approximately proportional to T. Therefore, approximately, the velocity v varies with temperature, as does q. If in an actual cell we would have to deal only with a single reaction, we would expect the rate of that reaction to increase ex-

ponentially with temperature, according to the Arrhenius equation, that is, to vary like $e^{-\varphi/T}$ (φ being a constant). As we have seen in chapter iii, q is actually a composite quantity, and therefore the relation $q \sim e^{-\varphi/T}$ can hold only approxi-

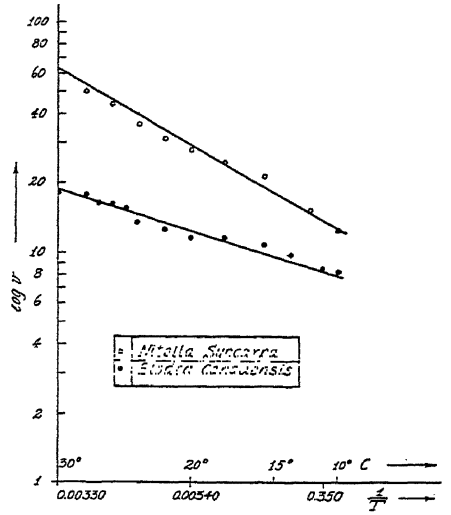


FIG. 26.—Logarithms of velocity of protoplasmic streaming plotted against the inverse of absolute temperature. Experimental data from C. Oppenheimer and L. Pincussen (1927), *Tabulae biologicae*, 4, 465 (Berlin: W. Junk). The relationship is approximately linear, as required by the theory.

mately. To the extent of this approximation we should, however, have $v \sim e^{-\varphi/T}$, or

$$\log v = \text{const} - \frac{\varphi}{T}. \qquad (6)$$

In other words, if we plot on semilogarithmic paper the logarithm of the velocity of streaming against the inverse tempera-

ture, we should have approximately a straight line with a negative slope. To what extent this relation is actually satisfied, is shown in Figure 26.

REFERENCES

1. L. V. Heilbrunn, *The Colloid Chemistry of Protoplasm* (Berlin: Gebr. Borntraeger, 1928).

2. Gale Young, *Bull. Math. Biophysics*, 1, 177, 1939.

3. *Tabulae biologicae*, 4, 465, 1927 (Berlin: W. Junk).

4. A. Gyemant, *Grundzüge der Kolloidphysik*, p. 55 (Braunschweig: Vieweg & Sohn, 1925).

CHAPTER VII

EXCITATION IN PERIPHERAL NERVES

In this chapter we shall discuss another rather general property of living cells, namely, their irritability. While, to some degree, irritability is found in almost all cells, the nerve cells exhibit it to a particularly high degree. A great many different physical and chemical factors, when acting on a limited region of an irritable cell, produce at that region a peculiar physicochemical state characterized by a local change in the metabolic rates, an increased general permeability, and an electronegativity relative to the nonaffected regions of the cell. This physicochemical state has the property of spreading to adjacent regions which are not directly affected by the external local physical or chemical factors. It thus propagates all over the cell. In the case of a nerve fiber, which is a highly elongated part of a cell body, this physicochemical change, which we call "excitation," propagates along the fiber, traveling over appreciable distances.

Another peculiarity of the excitatory state is that it usually develops suddenly, after the intensity of the external physical or chemical factors which induce it exceeds a rather definite threshold value.

The problem which confronts the mathematical biophysicist in this case is to derive the existence and properties of excitation from other basic biophysical properties of the cell, such as metabolic activity. In view of the great generality of excitatory phenomena, such an attempt is highly justified. Unfortunately, until now only a few general indications are available as to how phenomena of excitation may be reduced to other more general ones. The actual solution of the

problem is still wanting. This, however, should not detain us from a purely formal mathematical study of the properties of excitation, though we may not know what the true nature of excitation is. In the early development of physics, phenomena of heat, optics, and electromagnetism were treated as separate, rather unconnected, branches. Now we know that optical phenomena are merely a special case of electromagnetism. But even long before the physicists knew that, they were quite able to develop the mathematical theory of optical phenomena, though the true nature of light was unknown to them. And, when this true nature was revealed by the work of Maxwell, all the studies on mathematical optics were automatically incorporated into the general electromagnetic theory of light.

We shall follow the same procedure with respect to the excitation phenomena, pending their hoped-for reduction to more basic phenomena of cell biophysics.

Under physiological conditions different types of nerves require for their excitation different physical or chemical factors. Thus, the taste nerve fibers are excited by chemical substances. The acoustic nerve fibers and those of tactile sense respond to mechanical pressure changes, etc.

Under artificial conditions of the experiment we find, however, that there is one, so to say, universal factor for excitation, namely, the electric current. The protoplasm of the nerve cell through which any exciting current passes is an electrolyte. Therefore, the most obvious and immediate effect that the passage of a current may have upon the protoplasm is a displacement of various ions in one or another direction, depending on the sign of their electric charge. Such a displacement will result in local variations of concentrations of different ions. On the other hand, the classical work of J. Loeb and others has shown that the excitation state may be induced in many cells by treating them with solutions of ions, and that a certain threshold concentration of ions in the

solution must be exceeded in order to induce excitation. Considerations similar to these led W. Nernst,[1] in 1899, to assume that the excitation by electric currents is due to such changes in the concentrations of ions which, when they exceed a threshold value, induce excitation.

Suppose we let an electric current of some kind pass through a nerve. Then, if we can set up the differential equations which describe the motion of ions and the change of their concentrations under the influence of such a current, and if we can integrate these equations, we should be able to predict how long such a current would have to pass through the nerve in order to excite it.

The differential equations involved in such considerations are partial differential equations of a form whose integration sometimes presents very great difficulties. Setting up and solving a three-dimensional problem with approximately axial symmetry, as we have it in a nerve fiber, presents almost insuperable difficulties. Therefore, beginning with Nernst, different investigators[2] have resorted to exact solutions of highly oversimplified cases which had hardly any properties in common with the real cases. Yet some general results, agreeing with observations, were obtained. The situation was, to some extent, similar to the exact treatment involving the diffusion problem for a perfectly spherical cell, a problem discussed in mathematical biophysics, except that the unsatisfactory aspects of the procedure were even more apparent in the case of nerve excitation.

We have seen how, in the case of mathematical biophysics of the cell, the introduction of a rather drastic but powerful approximation method reduces problems of partial differential equations to those of ordinary ones, and problems of ordinary differential equations to problems of algebraic equations. A procedure similar to that seems indicated, therefore, in the case of nerve excitation.

If an electric current is flowing through the nerve fiber, then

the positive ions are driven to the cathode; the negative ones, to the anode. Let us first confine ourselves to the former. Although the actual concentration of the ions will vary from point to point in a complex way, the average concentration, $\bar{c}$, will increase at the cathode approximately at a rate proportional to the flowing current, I. The coefficient of proportionality, K, may be roughly evaluated from the average initial concentration c_0, the mobility of ions, and other electrochemical properties of the protoplasm and cell membrane. This particular problem of determining K in terms of other constants will not be discussed here.

Any excess of concentration will tend to decrease either because of a diffusion-like process or because of a chemical reaction which involves the destruction or removal of the ions. The rate of decrease will be roughly proportional to the amount of excess, $\bar{c} - c_0$. Denoting, here, the coefficient of proportionality by k, we thus arrive at the differential equation

$$\frac{d\bar{c}}{dt} = KI - k(\bar{c} - c_0)\,, \tag{1}$$

in which I is some function of the time t.

If, as said above, we assume that excitation occurs when the concentration $\bar{c}$ exceeds a critical value c^*, then, by introducing the new variables

$$\epsilon = \bar{c} - c_0; \qquad h = c^* - c_0\,, \tag{2}$$

we have

$$\frac{d\epsilon}{dt} = KI - k\epsilon\,, \tag{3}$$

with excitation occurring for

$$\epsilon \geqq h\,. \tag{4}$$

Equations (3) and (4) form the basis of H. A. Blair's theory of excitation,[3] which was proposed several years ago, although the argument given here and leading to its approach was not used.

In the absence of any current I, the quantity ϵ varies according to

$$\epsilon = \epsilon_0 e^{-kt}, \tag{5}$$

where ϵ_0 is the value which ϵ has at the moment $t = 0$. This shows that, when for a sufficiently long time the nerve has not been subject to any current, $\epsilon = 0$ and $\bar{c} = c_0$, as should be the case, physically. If, under these conditions, we now establish a sudden current of constant intensity I and count the time from the moment of the establishment of the current, then the quantity ϵ varies according to

$$\epsilon = \frac{KI}{k}(1 - e^{-kt}). \tag{6}$$

It tends asymptotically to the limiting value KI/k; and the initial slope of the ϵ, t line is equal to KI. Both the limiting value and the initial slope increase with increasing current I (*MB*, p. 151). If, in order to produce excitation, the value of ϵ must exceed the threshold h, then KI/k must exceed h; or, in other words, the current I must exceed a critical value $R = kh/K$. This value R is technically known as the "rheobase." Once the rheobase is exceeded, the quantity ϵ exceeds the threshold h after a time t, this time-interval decreasing as I increases. The relation between the intensity of the exciting constant current and the time which this current must take in order to produce excitation is obtained by equating the right-hand side of equation (6) to h and by solving this with respect to t. As has been shown by H. A. Blair, the relation thus obtained is in fair agreement with experimental data.[3]

One important conclusion is to be drawn from the comparison of the theoretical results with the experiments: the decrease of $\epsilon = \bar{c} - c_0$ cannot be due to any diffusion-like process. The order of magnitude of the coefficient k is about $10^2 - 10^3$ sec^{-1}. In other words, it must take about one-hundredth of a second or less to re-establish practically the initial concentration. As we have seen in chapter i, with plausible values for diffusion coefficients the time it takes for a diffusion process to reach a stationary state is of the order of a few seconds at least (cf. also *MB*, p. 12). Notice that equation (1) is essentially the same as equation (20) of chapter i, except that, instead of the accumulation of substance, owing to the rate of production q, we have an accumulation due to the current I. Hence, we should have $k = 1/\Lambda$. But with any plausible values for the constants involved in Λ, we cannot obtain such high values of k, and that discrepancy is found by comparing Blair's theory with experiments. And this holds also for the generalizations of Blair's theory, to be discussed later in this chapter. The only conclusion is that, while the ions are accumulated by the current, their concentration decreases, owing to some chemical reaction whose rate is proportional to $\bar{c} - c_0$. The initial finite concentration c_0 may be due to the fact that there are two reactions going in opposite directions—one producing, the other destroying, the particular type of ions. These two reactions would be in equilibrium for the particular value c_0 of the concentration.

Instead of a suddenly established constant current, we may consider any other kind of current, such as an exponentially varying current or an alternating current. This amounts to making the quantity I in equation (3) a definite function of time, integrating the equation, and again setting $\epsilon = h$. In this way a number of other results in more or less good agreement with experiments were obtained by H. A. Blair[3] (cf. also *MB*, chap. xvi).

There are, however, several difficulties of a fundamental nature which are not taken care of by H. A. Blair's theory. One of the most important is the behavior of a nerve at the anode of an electric circuit. We have hitherto confined ourselves to the cathode, to which positive ions are driven by the current. When a nerve is made to be a part of an electric circuit, and a current is suddenly established, excitation always occurs at the cathode. This indicates that the excitatory ions are positively charged. But the establishment of a constant current of indefinite duration usually results in a single excitation impulse, which may be observed either by a short twitch produced in a nerve-muscle preparation or by oscillographic recording of the propagation of the action potential accompanying the excitatory process. After the first impulse the nerve is not excited, though a constant current is passing through it. But if such a constant current is suddenly interrupted, then another excitatory impulse is observed, this time originating at the *anode* of the circuit. This phenomenon offers very great difficulties to any theory based on the assumption that excitation is induced by an excess of concentrations of some ions. If positive ions are excitatory, as seems evident, then at the anode their concentration decreases, and a sudden break of current results merely in the restoration of the initial concentration. One might think of some inertial effects which would make the concentration of the ions at the anode "swing" over the initial normal value. But the slowness of the diffusion processes makes the introduction of such a hypothesis physically quite untenable. Thus, until recently, the "break excitation" at the anode presented a stumbling-block to all theories.

In 1933 the author[4] suggested a natural generalization of H. A. Blair's theory, which solves the difficulty in a simple way. It is known that the protoplasm of any cell contains not one, but many, kinds of ions. Moreover, not all positive ions

produce excitation. Some of them inhibit the excitatory effects of other ions.[5] This suggests the study of a case where we have not one, but at least two, types of ions—one excitatory, the other inhibitory—both positively charged and therefore driven to the cathode. Denoting the concentration of the former by ϵ and of the latter by j, we assume that for each of them an equation of the form of the equation (1) holds* except that the constants involved are different for the two types of ions. Thus, we now have two equations:

$$\frac{d\epsilon}{dt} = KI - k(\epsilon - \epsilon_0) \tag{7}$$

and

$$\frac{dj}{dt} = MI - m(j - j_0)\,, \tag{8}$$

where again K, M, k, and m are constants and ϵ_0 and j_0 are the initial normal concentrations of ϵ and j. In line with our knowledge of the antagonism of ionic action,[5] we assume that excitation occurs when the ratio ϵ/j exceeds a critical value h. As far as single nerve fibers are concerned, we may, without loss of generality, put $h = 1$.† We thus have, as a condition of excitation,

$$\epsilon \geqq j\,. \tag{9}$$

Since in the absence of any current the nerve is not excited, we also have

$$\epsilon_0 < j_0\,. \tag{10}$$

* Note that the symbol ϵ stands here for a quantity different from that in equation (2).

† Cf. the following chapter, p. 132, concerning the cases when this restriction is not applicable.

Furthermore (*MB*, chap. xvii), we make the following assumption about the constants K, M, k, and m:

$$k \gg m\ ; \qquad K \gg M\ ; \qquad \frac{K}{k} \gtrsim \frac{M}{m}\ . \tag{11}$$

Figure 27 represents graphically what happens at the cathode and at the anode when a constant current, I, is suddenly

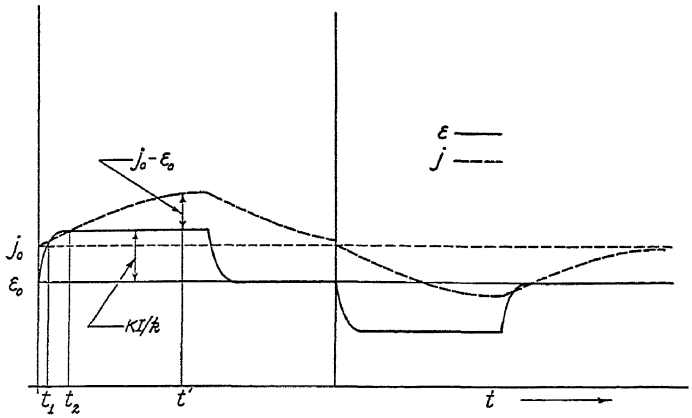


FIG. 27

closed or opened. Owing to inequalities (11), the initial slope of the ϵ-line is greater than that of the j-line. The quantity ϵ also approaches its limiting value more rapidly. As a result, ϵ exceeds j for a short time, during which excitation takes place at the cathode. Upon opening the current, the quantity j at the cathode decreases less rapidly than the quantity ϵ, because of $k > m$. Thus, in this case ϵ always remains below j, and no excitation occurs.

At the anode the variation of ϵ and j is represented by a pair of equations similar to equations (7) and (8) but with the sign of I reversed. Both ϵ and j decrease and for sufficiently strong currents may become negative, which is physically

impossible if they represent concentrations. The solution of the difficulty lies in the fact that equations (7) and (8) are only approximate and hold only within a limited range of currents. For that matter, the empirically found simple laws of excitation also hold only within a limited range of exciting currents. When the current at the anode is broken, both ϵ and j return to their initial values ϵ_0 and j_0, according to equations which are of the form of equations (5). Because of $k > m$, the quantity ϵ rises more rapidly, and again for a short time it exceeds the quantity j, so that excitation occurs.

Some other general facts, such as the failure of very slowly rising currents to excite, have been shown to follow naturally from this "two-factor" theory of excitation[4] (*MB*, chap. xvii). Other special cases, however, such as the time relations for excitation by alternating currents, were not studied in detail.

In 1935 A. V. Hill[6] proposed a somewhat different theory, which, however, shows much similarity to the above-mentioned one. A. V. Hill and his collaborators have subjected different consequences of that theory to thorough experimental tests and have found them in good agreement with experimental results.[7, 8]

A. V. Hill prefers to talk not of "ion concentrations" but rather of two factors, without specifying their physical nature. This has the advantage of a greater generality. For that matter, nothing would be changed in the author's theory if, instead of ionic concentrations, we would refer to ϵ as the "excitatory factor" and to j as the "inhibitory factor." We then would not even need to restrict ourselves to positive values of ϵ and j, as has been done on pages 109–10. A. V. Hill calls j the "accommodation factor."

As differential equations for the variation of the two factors, A. V. Hill chooses, in our notations, the following:

$$\frac{d\epsilon}{dt} = K'I - k'(\epsilon - \epsilon_0) \qquad (12)$$

and

$$\frac{dj}{dt} = M'(\epsilon - \epsilon_0) - m'(j - j_0)\,, \tag{13}$$

with relation (9) as condition of excitation. Moreover, Hill imposes a more restricted condition than (11), namely,

$$\frac{K'}{k'} = \frac{M'}{m'}\,. \tag{14}$$

Equation (12) is identical with our equation (7). However, equation (13) states that the rate of increase of the factor j is proportional not to the current I directly but to the excess $\epsilon - \epsilon_0$ of the excitatory factor, created by the current.

The general situation for "make" and "break" of a constant current is represented in Hill's theory by a diagram essentially similar to that of Figure 27. The only difference is that the j-curve is not a simple exponential one but begins with a zero slope and has an inflection point, as shown on Figure 28.

In view of the great similarity of the two theories, it seemed of interest to compare the detailed consequences of both. In attempting to do so, F. Offner[9] found that, as far as intensity-time relations for excitation with any kind of current are concerned, the two theories are entirely identical. Gale Young[10] proved subsequently that any other two-factor theory which is based on a pair of simultaneous differential equations, linear in the current I and in the factors ϵ and j, and with real characteristic roots, can be brought, by a suitable transformation, into the form given by equations (7) and (8). Thus, any experiments on excitation by electric currents which are in agreement with A. V. Hill's equations (12) and (13) are also in agreement with our equations (7) and (8). In-

asmuch as equations (7) and (8) are independent, while equations (12) and (13) are simultaneous, the mathematical analysis is somewhat simplified in the former case. As an illustration, we shall derive here from our equations, following A. Weinberg, the relations for excitation by alternating currents and for excitation at the anode—relations which have been studied theoretically and experimentally by A. V. Hill

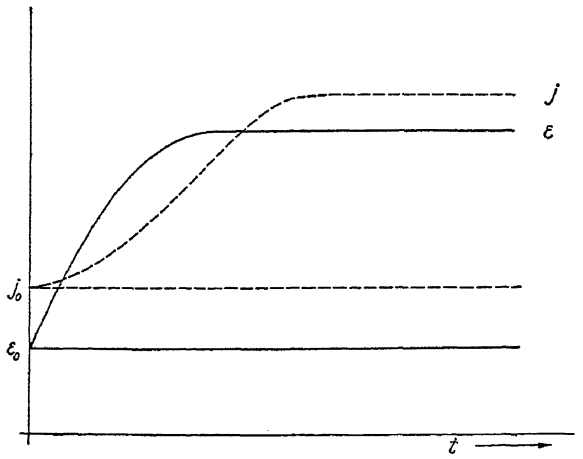


FIG. 28

and his collaborators. For a number of further details we must refer to the original paper of A. V. Hill and his school, covering a very extensive field. Although the fundamental idea of the two-factor theory had been suggested by the author a few years previously, to A. V. Hill and his school belongs all the credit for actually elaborating the theory in detail and subjecting it to rigid experimental tests.

Let the current I in equations (7) and (8) be of the form

$$I = I_0 \sin \omega t . \qquad (15)$$

Introducing equation (15) into equations (7) and (8), and integrating them by elementary methods (*MB*, p. 157), we find

$$\begin{aligned} \epsilon - \epsilon_0 &= \frac{I_0 K}{k^2 + \omega^2} (k \sin \omega t - \omega \cos \omega t) ; \\ j - j_0 &= \frac{I_0 M}{m^2 + \omega^2} (m \sin \omega t - \omega \cos \omega t) . \end{aligned} \tag{16}$$

Hence,

$$\left.\begin{aligned} \epsilon - j = I_0 \Bigg[\Bigg(&\frac{Kk}{k^2 + \omega^2} - \frac{Mm}{m^2 + \omega^2} \Bigg) \sin \omega t \\ &- \left(\frac{K\omega}{k^2 + \omega^2} - \frac{M\omega}{m^2 + \omega^2} \right) \cos \omega t \Bigg] + \epsilon_0 - j_0 . \end{aligned} \right\} \tag{17}$$

By a procedure similar to that used before (*MB*, p. 158), we find that the amplitude or the maximum value of $\epsilon - j$ is equal to

$$\left.\begin{aligned} &(\epsilon - j)_{\max} \\ &= I_0 \sqrt{\left(\frac{Kk}{k^2 + \omega^2} - \frac{Mm}{m^2 + \omega^2} \right)^2 + \left(\frac{K\omega}{k^2 + \omega^2} - \frac{M\omega}{m^2 + \omega^2} \right)^2} \\ &\qquad + \epsilon_0 - j_0 . \end{aligned} \right\} \tag{18}$$

After elementary rearrangements, using equation (14), which is merely a restricted form of the last equation (11), we find

$$(\epsilon - j)_{\max} = I_0 \frac{(K - M)\omega}{\sqrt{(k^2 + \omega^2)(m^2 + \omega^2)}} + \epsilon_0 - j_0 . \tag{19}$$

As we have seen in *MB*, page 167, equation (13), equations (7) and (8) lead to the following approximate expressions for

the rheobase R_0 of the cathodic excitation by a constant current

$$R_0 = \frac{(j_0 - \epsilon_0)k}{K - M} . \tag{20}$$

Hence,

$$j_0 - \epsilon_0 = \frac{R_0(K - M)}{k} . \tag{21}$$

Excitation by an alternating current will just occur when the amplitude I_0 of the current is high enough to make

$$(\epsilon - j)_{\max} = 0 . \tag{22}$$

Introducing equation (20) into equation (19) and the latter into equation (22), we find

$$\frac{I_0}{R_c} = \sqrt{\left(1 + \frac{\omega^2}{k^2}\right)\left(1 + \frac{m^2}{\omega^2}\right)} . \tag{23}$$

Equation (23) gives us the ratio of the liminal amplitude of an alternating sinusoidal current to the liminal value of a constant current in terms of the frequency ω. A. V. Hill has derived a similar relation on the basis of his theory, expressed by equations (12) and (13). In Hill's notations our k is denoted by $1/k$, our m is denoted by $1/\lambda$, and ω is equal to $2\pi n_1$. Moreover, Hill uses I_0 for our R_c and I for our I_0. With this change of notations equation (23) becomes identical with equation (1) of A. V. Hill, B. Katz, and D. Solandt.[7] Figure 29 shows a comparison of the equation with experimental data, taken from a paper by these authors.

We now shall derive from equations (7) and (8) the rela-

tion between the duration of a constant current and the minimum intensity necessary to produce an excitation at the anode upon breaking the current. If $\bar{t}$ represents the duration of the current I from the moment of the "make" to the mo-

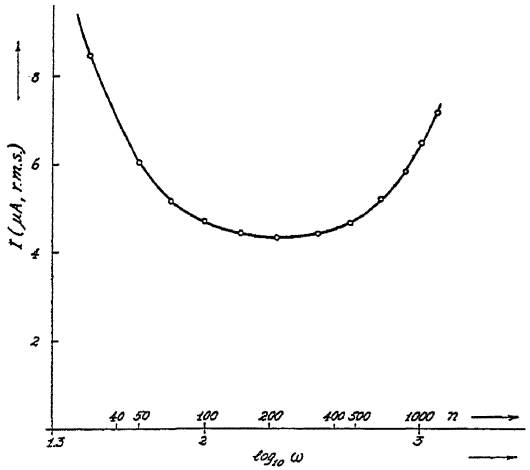


FIG. 29.—Excitation by alternating current of frequency ω. Threshold current I plotted against log $_{10}\omega$. The curve is the theoretical one, derived from equation (23). Circles represent observed values. From A. V. Hill, B. Katz, and D. Y. Solandt, *Proc. Roy. Soc. London*, B, 121, 113, 1936.

ment of the "break," then, according to equation (6) and the remark about behavior at the anode made on page 109, at $t = \bar{t}$ the quantities $\epsilon - \epsilon_0$ and $j - j_0$ are given by

$$\begin{aligned} \epsilon - \epsilon_0 &= -\frac{KI}{k}\left(1 - e^{-k\bar{t}}\right); \\ j - j_0 &= -\frac{MI}{m}\left(1 - e^{-m\bar{t}}\right). \end{aligned} \tag{24}$$

From the moment $t = \bar{t}$ on, these values vary according to equations of the form of equation (5) with the right-hand sides of (24) as initial values. Hence,

$$\left.\begin{aligned} \epsilon - \epsilon_0 &= -\frac{KI}{k}\left(1 - e^{-k\bar{t}}\right)e^{-kt}; \\ j - j_0 &= -\frac{MI}{m}\left(1 - e^{-m\bar{t}}\right)e^{-mt}. \end{aligned}\right\} \quad (25)$$

As we have seen in *MB*, chapter vii, when excitation occurs just at threshold we have $d\epsilon/dt = dj/dt$. Introducing equations (25) into this requirement and solving the resulting equation for t, we find for the time of excitation after rearrangements, again remembering equation (14),

$$t = \frac{1}{m - k} \log \frac{m}{k} \frac{e^{-m\bar{t}} -}{e^{-k\bar{t}} -} \quad (26)$$

On the other hand, at the moment of excitation we have $\epsilon = j$. Introducing expression (26) for t into equations (25), and the latter equations into $\epsilon = j$, we find, after some rearrangements,

$$\epsilon_0 - j_0 = \frac{KI}{k} \frac{\left(1 - e^{-k\bar{t}}\right)^{m/(m-k)}}{\left(1 - e^{-m\bar{t}}\right)^{k/(m-k)}} \left(\frac{k}{m}\right)^{m/(m-k)} \left(\frac{m}{k} - 1\right). \quad (27)$$

Since $m \ll k$, $(m/k) - 1 \sim -1$; and therefore equation (27) gives approximately

$$\frac{KI}{k(j_0 - \epsilon_0)} = \frac{\left(1 - e^{-k\bar{t}}\right)^{1/[(k/m)-1]}}{\left(1 - e^{-m\bar{t}}\right)^{1/[1-(m/k)]}} \left(\frac{k}{m}\right)^{1/[(k/m)-1]}$$

Again making use of equation (21) and of inequalities (11), we finally obtain

$$\frac{I}{R_c} = \frac{\left(1 - e^{-k\bar{t}}\right)^{1/[(k/m)-1]}}{\left(1 - e^{-m\bar{t}}\right)^{1/[1-(m/k)]}} \left(\frac{k}{m}\right)^{1/[(k/m)-1]} \quad (28)$$

Equation (28), except for differences in notations, is again identical with the corresponding equation deduced by A. V. Hill from his equations (12) and (13). Figure 30 shows a comparison of that equation with experimental data.

A problem closely connected with that of excitation is the problem of conduction of the nervous impulse. It has been rather generally accepted that the cause of the spread of

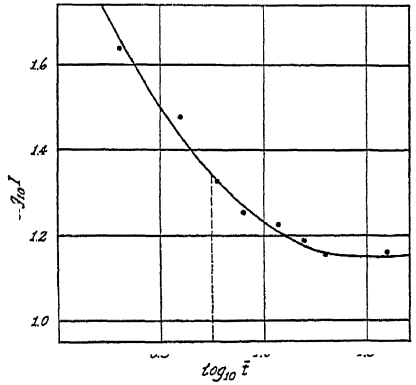


FIG. 30.—"Break" excitation at end of constant current pulse. I represents current threshold for excitation at pulse duration t. Points are experimental values. The curve is the theoretical one, derived from equation (28). From D. Y. Solandt, *Proc. Roy. Soc. London*, B, **120**, 400, 1936.

excitation lies in the local electronegativity of the excited region. This electronegativity with respect to adjacent non-excited regions results in an electromotive force which sets up local electric currents. These currents act in such a way as to excite directly the regions immediately adjacent to the already excited one.

It follows, from this picture, that the mechanism of propagation of the excitation along the nerve will depend largely on the mechanism of excitation by electric currents. It also depends on the electrical properties of the nerve fiber. Assuming H. A. Blair's theory of excitation and neglecting the

distributed capacity of the nerve fiber, the problem of conduction can be treated exactly as has been shown by the author;[11] it requires the solution of a special type of functional equation. For the two-factor theory of excitation in the author's form, the problem still can be solved exactly,[12] though the exact expressions are rather cumbersome, and considerable simplifications are obtained by some approximations. Whereas Blair's theory leads to a strictly constant velocity of conduction, the two-factor theory requires that the rate of conduction should vary very slightly in the neighborhood of the originally excited region, but within about a millimeter of this region, there will be practically a constant value. The experimental detection of such slight variations would present some difficulties.

Alvin Weinberg[13] has studied the problem for the case of a nonnegligible capacity of the nerve. In this case only an approximate solution has been obtained so far.

All the above-mentioned theories lead to expressions for the velocity of propagation in terms of a number of other constants of the nerve fiber, such as resistance of myelin sheath, rheobase, etc. Most of these constants have, unfortunately, not yet been determined with any degree of accuracy. Therefore, while we may say that the expressions obtained agree in a general way with observations, a really quantitative and exact verification is still wanting. At any rate, these mathematical studies definitely suggest new experimental problems.

A theory of nerve conduction, based on an approach somewhat different from the above-mentioned ones, has been developed in an interesting paper by W. A. H. Rushton.[14]

REFERENCES

1. W. Nernst, *Göit. Nachr.; Math.-Phys. Klasse*, p. 104 (1899); *Pflüg. Arch.*, **122**, 275, 1903.

2. K. Umrath, *Biologia generalis*, 1, 396, 1924.

3. H. A. Blair, *Jour. Gen. Physiol.*, 15, 709, 731, 1932; 18, 755, 1935; *Amer. Jour. Physiol.*, 111, 515, 1935; 112, 277, 1935; *Jour. Cell. and Comp. Physiol.*, 6, 291, 1935; *Cold Spring Harbor Symposia on Quantitative Biology*, 4, 63, 1936.

4. N. Rashevsky, *Protoplasma*, 20, 42, 1933; *Cold Spring Harbor Symposia on Quantitative Biology*, 4, 90, 1936; *MB*, chap. xvii.

5. E. Gellhorn, *Lehrbuch der allgemeinen Physiologie* (Leipzig: G. Thieme, 1931); W. M. Bayliss, *Principles of General Physiology* (New York: Longmans, Green & Co., 1927).

6. A. V. Hill, *Proc. Roy. Soc. London*, B, 119, 305, 440, 1936.

7. A. V. Hill, B. Katz, and D. Y. Solandt, *Proc. Roy. Soc. London*, B, 121, 74, 1936.

8. D. Y. Solandt, *Proc. Roy. Soc. London*, B, 120, 389, 1936.

9. F. Offner, *Jour. Gen. Physiol.*, 21, 89, 1937.

10. G. Young, *Psychometrika*, 2, 103, 1937.

11. N. Rashevsky, *Physics*, 4, 341, 1933; *MB*, chaps. xix and xx.

12. N. Rashevsky, *Physics*, 6, 308, 1935; *MB*, chap. xxi.

13. A. Weinberg, *Jour. Applied Physics*, 10, 128, 1939.

14. W. A. H. Rushton, *Proc. Roy. Soc. London*, B, 124, 210, 1937.

CHAPTER VIII

SOME REMARKS ON THE PHYSICOMATHEMATICAL ASPECTS OF CENTRAL EXCITATION AND INHIBITION

In the third part of our book *Mathematical Biophysics* we attempted to develop a systematic abstract mathematical theory of the functions of the central nervous system. The fundamental idea was to postulate a few mathematically definite laws of interaction between two adjacent neuroelements and then to consider to what consequences such laws of interaction lead when applied to different geometrical arrangements of the interacting elements. In other words, the attempt was made to reduce the tremendous complexity of functions of the central nervous system to the complexity of its structure, keeping the fundamental dynamic processes as simple as possible. A different line of approach is also possible, namely, to consider theoretically the structure of the central nervous system, with its tremendous number of neuroelements, as quasi-homogeneous and to try to account for the enormous complexity of its functions by postulating correspondingly complex dynamical laws of interaction between the individual elements.

It appears to us that the first point of view may be considered as indirectly better justified by actual observation. As far as experience goes, there does not seem to exist any too pronounced fundamental difference between the elementary physiological properties of the neuron of the frog and those of the neuron of man. One would therefore naturally tend to explain the enormous difference in the functions of the brain of the frog and that of man by the equally enormous differ-

ence in the complexity of the structure of their brains. Theoretically, also, the first approach seems to be the more promising one. In some earlier publications[1] we have attempted to develop the second point of view. While some rather interesting general conclusions were thus obtained, nothing of a quantitative nature seems to be gained without a large number of additional and rather disconnected assumptions. The interesting speculations of W. Köhler,[2] which implicitly assume the second point of view, also do not materially contribute to a quantitative mathematical theory of the brain. On the other hand, the first point of view, as developed in *MB*, not only has led to a rather natural systematic mathematical theory of the brain but has also shown its empirical usefulness by leading to quantitative relations that are found to be verified experimentally. We refer to H. D. Landahl's work on psychophysical discrimination and on reaction times, to A. S. Householder's study of the sensory discrimination, and to N. Rashevsky's work on the aesthetic measure of geometrical patterns, to be discussed in subsequent chapters.

While the primary object of the developments given in *MB* was to work out an abstract physicomathematical theory of the brain and to demonstrate the applicability, in principle, of the mathematics to this field, the above-mentioned work of Landahl, Householder, and Rashevsky does indicate that something more than a mere illustration is contained in our fundamental postulates. In choosing these postulates, we were guided by some general experimental evidence. On the other hand, however, the possible connection of these postulates with some other empirical evidence may not be so clear. We shall therefore analyze such a connection somewhat more elaborately.

Taking a suggestion from the well-known results of E. D. Adrian and his school, we first assume that under physiological conditions a constant peripheral stimulus of any nature

results in a sequence of impulses in each single afferent nerve fiber. The intensity I of each impulse is assumed to be constant, thus preserving the "all-or-none" law, while the frequency ν of the impulses is taken to be an increasing function of the intensity S of the stimulus. For small values of S the frequency ν is considered as approximately linear:

$$\nu = a(S - h)\,, \tag{1}$$

where h is the threshold below which a stimulus does not elicit any excitation and where a is a constant. Equation (1) is understood to be only an approximation to some more general type of equation, in which ν tends to a constant value with increasing S, as must be the case from considerations of the refractory state of the fiber (*MB*). Such a more general equation may be, for instance (*MB*, p. 218),

$$\nu = \frac{1}{\theta}\left[1 - e^{-a\theta(S-h)}\right],$$

where ϵ is the refractory period; or, as another possibility,

$$\nu = ah \log \frac{\dfrac{S}{h}}{1 + \delta \dfrac{S}{h}}\,; \qquad \delta \ll 1\,, \tag{2}$$

which, for small values of S/h, reduces to

$$\nu = ah \log \frac{S}{h}\,.$$

Defining the intensity, E, of excitation of a fiber by

$$E = I\nu\,, \tag{3}$$

which, together with (1), gives

$$E = aI(S - h)\,, \tag{4}$$

we now establish the fundamental equations of interaction between adjacent neuroelements.

The axon end of every neuron is assumed to produce two factors—one excitatory, ϵ; and the other inhibitory, j. It must be *emphasized* that we do not consider these factors as substances. They may be such, but they may represent any kind of physicochemical state having the property of exciting or inhibiting the neuron upon which it acts. We assume that the quantity $\epsilon - j$ affects any neuron in the same way as does a natural stimulus; that is, ϵ and j, produced by the ith neuron at the synapse with a kth neuron, result in an excitation of the latter with an intensity E_k, given, in analogy to (4), by

$$E_k = a_k I_k(\epsilon - j - h_k)\,. \tag{5}$$

Excitation of the kth neuron occurs only if $\epsilon - j > h_k$.

The equations necessary to complete our set of postulates specify the production of ϵ and j by a fiber. They are

$$\begin{aligned} \frac{d\epsilon}{dt} &= AE - a\epsilon\,; \\ \frac{dj}{dt} &= BE - bj\,. \end{aligned} \tag{6}$$

Equations (1)–(6) are sufficient to describe quantitatively any situation arising from the interaction of any number of neurons arranged according to a definite geometrical pattern. Provided we can master the purely mathematical complexity of any such problem, we can proceed to the consideration of various forms of neuronic circuits and their functions.

Equations (6) are formally identical with equations (7) and (8) of the preceding chapter. When a stimulus S, of constant intensity, is suddenly established and kept indefinitely, then, according to equation (4) or to a corresponding, more general one, the result is the sudden establishment of a constant E. And following this, there is, as in chapter vii, an exponential increase of ϵ and j, both tending asymptotically to limiting values, corresponding to AE/a and to BE/b. The larger the a or b, the more rapidly these limiting values are approached.

We do not need to make any definite assumptions about the constants A, B, a, and b in equation (6). The following cases, however, present special interest.

If

$$A > B\,; \qquad a > b\,; \qquad \frac{A}{a} > \frac{B}{b} \tag{7}$$

then for a constant stimulus of indefinite duration the fiber always exerts an excitating action on the other fibers with which it synapses, with ϵ being produced in excess of j. If, on the contrary,

$$A < B \qquad a < b\,; \qquad \frac{A}{a} < \frac{B}{b} \tag{8}$$

then the fiber is an inhibitory one (MB).

Whenever the constants satisfy the relations (7), so that the nerve is an exciting one, we shall attach to those constants the index e. We shall have, in this case, A_e, B_e, a_e, and b_e. Whenever the constants satisfy relations (8), we shall mark them A_i, B_i, a_i, and b_i. A neuron producing only ϵ shall be called "purely excitatory"; one producing only j, "purely inhibitory." While, in a sense, we thus assume the existence of special inhibitory fibers, an assumption which does not meet

at present with general recognition (see J. C. Eccles,[3] R. Lorente de Nò,[3] and H. S. Gasser[4]), yet, at the same time, according to this theory, an inhibitory fiber differs from an excitatory one only quantitatively, the two being otherwise fundamentally identical in their general behavior. Other relations between the constants A, B, a, and b lead, moreover, to situations in which one and the same fiber may act either as an excitatory or as an inhibitory one, depending on the time relations of the stimuli (MB).

We may also have cases intermediate between those represented by relations (7) and (8); namely, we may have

$$\begin{aligned} &A < B \quad \text{with} \quad a < b \quad \text{and} \quad \frac{A}{a} > \frac{B}{b}\,; \\ &A > B \quad \text{with} \quad a > b \quad \text{and} \quad \frac{A}{a} < \frac{B}{b}\,. \end{aligned} \tag{9}$$

The first case gives for a constant, continuous stimulus a brief inhibition, followed by a lasting excitation (MB, p. 224). The second case gives, for the same conditions, a brief excitation, followed by a lasting inhibition.

While, as has been remarked, the above-mentioned postulates lead, in a number of cases, to consequences which are in *quantitative* agreement with experimental data, and while such an agreement may be considered as justifying them, to some extent, still, at first sight, these postulates do not fit quite well into the picture made by neurophysiologists of the basis of some electrophysiological and other experiments.

First of all, the situation as represented by our equations (1)–(6) is undoubtedly highly oversimplified, as compared with the actual situation. The wealth of observations reviewed, for example, by Fulton or by Eccles[3] reveal a maze of complexity in the interaction of central neurons. While, according to our equations, the elementary process of excitation

and conduction at a synapse appears to be a rather simple one, we know that it is actually decomposable into a rather large number of factors.

It must, however, be kept in mind that, from the point of view of the theory of interaction of the very large number of neurons which constitute the central nervous system, the *details* of the mechanism of interaction of two neurons may be quite irrelevant as long as we have a quantitative expression which describes the *net result* of this interaction. And, while the details may be very complex, these net results may be described by relatively simple equations. An example from physics may best illustrate the point. In the theory of heat conduction it is *assumed*, on the basis of some generalizations of experimental observations, that the heat transport in a given direction is proportional to the gradient of temperature T in that direction. Thus, choosing the direction of flow Q as that of the x-axis, we have

$$Q = -\kappa \frac{dT}{dx}, \tag{10}$$

where κ is the thermal conductivity. On this simple relation is based a mathematically most complete and practically very useful theory of heat conduction in bodies of various shapes. Yet the actual mechanism of heat transport is a very complex one, determined by the detail of the thermal motion of the molecules. The constant κ is a constant only to a first approximation, and even then it resolves itself into a number of factors, expressed in terms of gas-kinetic parameters. The complete and exact expression for the thermal conductivity κ is not available even now, though we have a pretty good insight at present into the mechanism of heat transfer. At the time of Fourier, who founded the mathematical theory of heat conduction, practically nothing was known of the com-

plexities of the kinetic theory. Yet, without waiting for the detailed mechanism which underlies equation (10) to be worked out, the theory of heat flow has reached a high degree of theoretical and practical perfection.

Similarly, the simple equations (5) and (6), which form the essence of our theory of central excitation and inhibition, may be used for a development of the theory without waiting for a detailed physiological interpretation of the underlying mechanism. Unquestionably, these equations are very crude and approximate, but so is equation (10) of the heat transfer. Yet for many practical purposes they may be accurate enough, as the above-mentioned work of Landahl, Householder, and Rashevsky indicates. The factors ϵ and j may each be of a rather composite nature, just as is the coefficient κ in equation (10). A physicochemical interpretation of these factors must be hoped for in the future. At present we may justifiably proceed on the more formal level. From this point of view it may be proper to call the developments of the third part of *MB* and such developments in subsequent publications "mathematical neurology" rather than "mathematical biophysics of the central nervous system," to emphasize the lack, *at the present stage*, of any physiological and physical interpretation.

It has been emphasized recently, on many sides, that the inhibitory effects are more likely to be due to an interaction of specially arranged excitatory mechanisms than to specific inhibitory impulses. Some such mechanisms have been suggested, for instance, by H. S. Gasser.[4] The net result of the functioning of such a scheme or any other similar one is the creation of an inhibitory state, which gradually disappears. If the rate of establishment of this inhibitory state is approximately proportional to the intensity of excitation of the neuron responsible for setting in action the whole mechanism, and if its decay, in the absence of disturbances, is ap-

proximately exponential, things will occur as if they were described by equations (5) and (6). The j-factor in this case may stand for a very complex excitatory state of a rather complicated anatomical structure. All that matters is the end result.

It must be emphasized that the element of "as if" enters into *every physical problem described by differential equations*. A differential equation deals with infinitesimals at a mathematical point, and, while it may be suggested by experiments, it is actually beyond any *direct* experimental verification. In many cases the most important differential equations of physics have only a statistical meaning and are known to break down for very small regions. Such is the case, for instance, with the differential equations of hydrodynamics and aerodynamics. Concepts like pressure, density, etc., which enter as continuous functions into these equations, have a meaning only for volumes which contain a very large number of molecules. For a volume of 10^{-18} cm^3 these concepts break down, and the differential equations do not hold at all. Yet, with very great accuracy, things actually happen *as if* the differential equations would hold down to infinitesimal regions. A similar interpretation is possible for our fundamental postulates (1)–(6). In the future they must be reduced to the details of physicochemical mechanism, just as equation (10) is now being reduced to kinetic interpretations. There is no doubt that a more detailed interpretation of our postulates will result in their modification and improvement, and a study of possible modifications is definitely indicated. In the meantime, however, we can and should proceed with the development of the theory based on these postulates.

From these general remarks we shall now pass to the discussion of a few specific questions of the relation of our postulates to some current neurophysiological views.

The intensity E of excitation of a fiber, as defined by (3),

can have only a statistical significance. When we are dealing with only one, or a very few, irregularly spaced impulses, such as are used in electrophysiological experiments, the concept of frequency, ν, loses its meaning.

Considering continuous stimuli, in which case (3) can be used, we find (*MB*, chap. xxiv) that the synaptic delay between *two* neurons decreases with increasing intensity of the stimulus. In the general case the synaptic delay is *finite* at threshold, decreasing to zero with increasing S. This latter decrease, however, will occur only if we consider equation (1) as holding for any value of S. If, for higher values of S, we adopt equation (2), then, as is readily seen from the argument in *MB*, increasing S decreases the synaptic delay only down to a finite value. Thus the synaptic delay varies within fixed limits; and by a proper choice of constants we may make those limits sufficiently narrow so as to be in agreement with the experimental findings of R. Lorente de Nò.

Let us now look at the picture from a different angle. Suppose that the transmission at the synapse is due to the action current of the presynaptic impulse. Two possibilities must then be considered: (*a*) the intensity of a single "all-or-none" impulse is subliminal; (*b*) the intensity of a single impulse is superliminal.

In case (*a*) a volley of impulses of frequency ν will produce a pulsating excitation current at the synapse which will act much as an *average* equivalent direct current of constant intensity would. Considering, for simplicity, the shape of the action current of an impulse as rectangular, we can calculate the excitation time of the neuron at the synapse by using the equations developed in *MB*, pages 220–26. If we assume that the law of excitation of the neurons is the same as that of peripheral nerves, based on the Hill-Rashevsky two-factor theory, we shall find for the synaptic delay formally the same expression as before.

In case (*b*), however, the synaptic delay between *two* neurons will be determined solely by the intensity of the individual impulse. It will be quite constant for a given pair of neurons and will not vary with either E or S. If, however, several fibers of different thresholds form synapses with the body of the same neuron, then increasing the intensity of the stimulus increases the number of excited fibers, and the body of the neuron is acted upon by the sum of several impulses, with a corresponding reduction of the synaptic delay. This seems to represent the case studied by R. Lorente de Nò.[3] While his experiments give strong evidence for this interpretation, there is no reason known, as yet, to assume that case (*a*) may not also be present in some parts of the central nervous system. R. Lorente de Nò's observations refer only to motoneurons and do not preclude a variability of the synaptic delay between *two* neurons of some other types, especially because the two cases differ only in the quantitative values of the parameters of the neurons.

Considering, now, the inhibitory effects, we may well assume that the inhibition is merely the result of a refractoriness produced by proper excitation. Let, for instance, a particular case be characterized by a very large a (Eq. [1]), so that even for very weak stimuli, S, the frequency, ν, will be so high that the successive impulses will fall within an interval of time smaller than the refractory period of another fiber with which the first one synapses. Then for average stimuli S, applied to the first fiber, the second fiber will be maintained continuously in a refractory state and thus may be considered as inhibited. It would be only natural to assume that, in general, the inhibitory state of the second fiber, as measured by the increase of the threshold, will be the stronger the higher the frequency ν of the first, "inhibiting" fiber. At the same time, it is not likely that this state of inhibition, or, if we prefer, of "relative refractiveness," will be established

at once, with the first impulse of the first fiber. If this state develops gradually at a rate approximately proportional to the frequency ν of the first fiber, and if, when left to itself, it decreases exponentially with respect to time, then, mathematically speaking, this refractory state will have, approximately, all the characteristics of our j-factor, as described by equation (5) and (6*b*). It must be remarked that equation (5) may be written thus:

$$E_\kappa = \alpha_\kappa I_\kappa[\epsilon - (h_\kappa + j)], \tag{11}$$

with

$$\epsilon > h_\kappa + j \tag{12}$$

as a condition of excitation. In this form the equation merely says that the effect of j is to increase the threshold of the other neuron.

If the first fiber synapses with a fiber which has a very short refractory period, then the former will excite the latter, producing in it a continuous volley of impulses of frequency ν. Thus, according to this conception, the same fiber may act as an excitatory or as an inhibitory one on another fiber, depending on the relative values of the physical constants of the two fibers.

Our equations (1)–(6) have been patterned very much along the lines suggested by the two-factor theory of peripheral excitation. In our original form[5] we called the factors "excitatory" and "inhibitory." In the slightly modified but essentially equivalent form of the theory proposed later by A. V. Hill[6] our inhibitory factor is called "accommodation." The name does not seem to matter much, since the effects of the increase of accommodation are opposite to those resulting from the increase of excitation, and so accommodation is essentially an inhibitory factor. The experimental evidence ob-

tained by A. V. Hill and his school has put the two-factor theory on a firm empirical basis. It is, therefore, quite in keeping with the prevailing tendency to reduce central excitation and inhibition to laws of peripheral nerves to adopt here, also, some sort of a "two-factor" theory, as we have done.

One might even go farther and identify the central ϵ and j with the peripheral ones. We may assume that ϵ and j produced at the end of the axon act directly on the adjacent neuron. In the theory of peripheral excitation we have considered (cf. *MB*) as a condition of excitation

$$\frac{\epsilon}{j} \gtrless h,^*$$

and for simplicity have chosen the units so as to make $h = 1$, which reduces (12) to

$$\epsilon \gtrless j.$$

This is admissible as far as a single fiber is concerned. But, in considering several fibers, we must take into account the possibility that h varies from fiber to fiber, in which case we cannot consider it as always equal to 1. If a fiber with a given h synapses with another, which has an $h' > h$, then while $\epsilon = hj$ in the first fiber produces excitation, it fails to do so in the second one, because there we must have $\epsilon = h'j > hj$. Inasmuch as, in the two-factor theory of peripheral excitation, ϵ first exceeds j and then again drops below it, thus creating a state of inhibition which for properly chosen constants may be stronger than the initial $(j_0 - \epsilon_0)$ (*MB*), several interesting possibilities are open here.

* The h in (12) is not to be confused with the h in equations (1) and (5) (cf. *MB*, chaps. xvii and xxii).

REFERENCES

1. N. Rashevsky, *Jour. Gen. Psychol.*, **5**, 368, 1931; **13**, 82, 1935.

2. W. Köhler, *Gestalt Psychology* (New York: Liveright, 1929).

3. J. C. Eccles, *Physiol. Rev.*, **17**, 538, 1937, and *Ergebnisse d. Physiol.*, **38**, 339, 1936; R. Lorente de Nò, *Jour. Neurophysiol.*, **1**, 187, 195, 207, 1938; J. F. Fulton, *Physiology of the Nervous System* (New York: Oxford University Press, 1938).

4. H. S. Gasser, *Volume Jubilaire publie en l'honneur du Prof. J. Demoor*, p. 218 (Liège, 1937).

5. N. Rashevsky, *Protoplasma*, **20**, 42, 1933, and *Cold Spring Harbor Symposia on Quantitative Biology*, **4**, 90, 1936.

6. A. V. Hill, *Proc. Roy. Soc. London*, B, **119**, 305, 440, 1936.

CHAPTER IX

MATHEMATICAL BIOPHYSICS OF SOME SIMPLE NEUROLOGICAL STRUCTURES: APPLICATIONS TO REACTION TIMES

Having established in the previous chapter the fundamental equations which govern the interaction between neurons, we shall now follow the procedure clearly indicated. We must systematically study different, more or less complex, geometrical arrangements of neurons and mathematically derive their properties. We shall start with the simplest possible cases and then gradually complicate the picture.

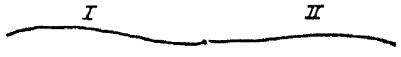


FIG. 31

The simplest possible structure is that of a neuron with a peripheral fiber synapsing with a neuron of higher order. As we have seen in the previous chapter, we do not need to enter into the considerations of the detailed structure of the synapse in order to apply our fundamental equations. We shall therefore represent two synapsing neurons schematically, as shown in Figure 31, simply by the adjacent lines. In MB we indicated by arrows, placed at the synapse, the direction of transmission. Inasmuch as a forklike scheme is used in neurological diagrams to represent schematically the branchings of the axon at the synapse and a point is used to represent schematically the body of the higher-order neuron, our former notation appeared somewhat confusing to the neurologists,[1] for in the classical neurological notations the nervous impulse

travels from the fork of one neuron to the point of the other, that is, in the direction opposite to that assumed in our notations, since a fork can be looked upon as an inverted arrow. Therefore, we shall now place the arrows which indicate the direction of transmission in the middle of the lines which represent the neurons, thus avoiding confusion. It must be noted, however, that some neurologists use the same convention as we used in *MB*.[2]

We may now consider various stimuli, S, which are different with regard to their variation with respect to time, as being applied to the afferent end of the neuron I (Fig. 31). We shall confine ourselves here to the consideration of a stimulus of constant intensity S_1, applied suddenly and kept indefinitely.

In this case neuron I will be subject to a suddenly established intensity of excitation E_1, connected to S_1 by one of the equations discussed in chapter viii. From the moment the excitation arrives at the synapse, the two factors ϵ and j will vary according to

$$\epsilon = \frac{A_e E_1}{a_e}(1 - e^{-a_e t}) ; \qquad j = \frac{B_e E_1}{b_e}(1 - e^{-b_e t}) . \qquad (1)$$

As soon as $\epsilon - j$ reaches or exceeds the threshold h_2 of the neuron II (Fig. 31), the latter becomes excited, and the excitation is transmitted farther along it. The time which elapses between the moment of arrival of the excitation along the fiber I at the synapse and the moment when the neuron II becomes excited—in other words, the synaptic delay—is obtained by solving, with respect to t, the equation $\epsilon - j = h_2$, after introducing into the latter the values for ϵ and j from equations (1). The details of the calculations are given in *MB*, chapter xxii. The relations become particularly simple when B_e is very small—in other words, when neuron

I produces very little of the inhibitory factor j, so that the latter may be neglected, as compared with ϵ. In this case the synaptic delay is given by (*MB*, p. 220, Eqs. [12] ff.)

$$t_s = \frac{1}{a_e} \log \frac{A_e E_1}{A_e E_1 - a_e h_2} . \tag{2}$$

The relation for the general case is more complicated, though similar in general, to the foregoing (*MB*, chap. xxii).

Equation (2) shows that the stronger the intensity of excitation E_1 the shorter the synaptic delay t_s. Since E_1 increases with increasing intensity S_1 of the peripheral stimulus, therefore the stronger the stimulus S_1 the shorter should be the synaptic delay t_s. To obtain a relation between t_s and S_1, we must use one of the equations of the previous chapter. Assuming relation (2) of chapter viii to hold, we find, approximately, for not too large values of S_1,

$$t_s = \frac{1}{a_e} \log \frac{A_e I_1 a_1 h_1 \log S_1 - A_e I_1 a_1 h_1 \log h_1}{A_e I_1 a_1 h_1 \log S_1 - A_e I_1 a_1 h_1 \log h_1 - a_e h_2} , \tag{3}$$

in which the subscript 1 indicates that the constants refer to the neuron I.

Putting

$$H_1 = \log h_1 ; \qquad H_2 = \frac{a_e h_2}{A_e I_1 a_1 h_1} ,$$

we obtain equation (3) in the form

$$t_s = \frac{1}{a_e} \log \frac{\log S_1 - H_1}{\log S_1 - H_1 - H_2} . \tag{4}$$

Using the approximate expression (4) of chapter viii, we find for t_s an expression of the form (*MB*, chap. xxii, Eq. [13])

$$t_s = \frac{1}{a_e} \log \frac{A_e a_1 I_1 (S_1 - h_1)}{A_e a_1 I_1 (S_1 - h_1) - a_e h_2} . \tag{5}$$

Consider a chain of neurons transmitting the excitation successively to each other, and remember that the velocity of propagation along each individual link of the chain is inde-

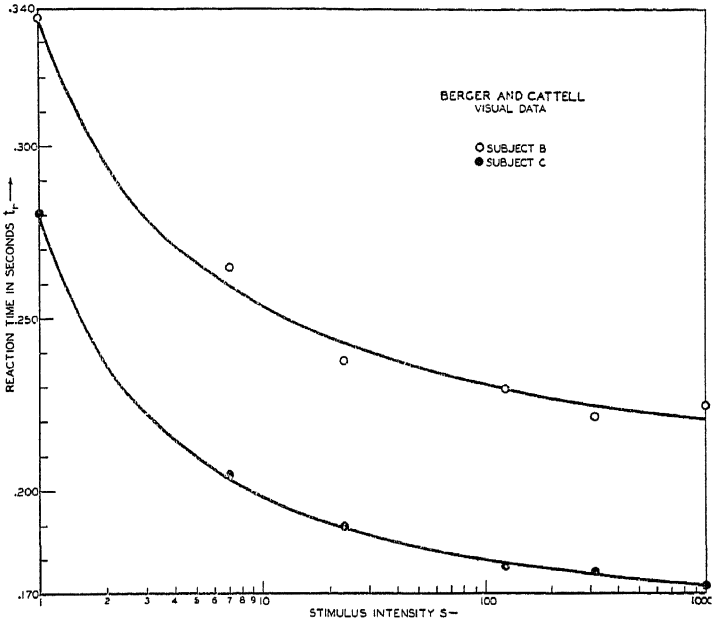


FIG. 32.—Reaction time for visual stimuli, plotted against stimulus intensity. The curves represent equation (4). The circles represent experimental values by G. O. Berger and J. McK. Cattell for two different subjects. The following values for the constants in equation (4) were used. Subject B: $1/a_e = 0.413$ sec; $H_1 = -1.96$; $H_2 = 0.57$. Subject C: $1/a_e = 0.279$ sec; $H_1 = -1.57$; $H_2 = 0.57$. From H. D. Landahl.[3]

pendent of the intensity of excitation. In that case a *qualitatively* similar relation between the intensity of the peripheral stimulus and the total time of transmission along the chain must hold, that is, the stronger the peripheral stimulus the shorter the total time of transmission along the chain. We

cannot expect equations (4) and (5), which were derived for a single synapse, to hold *quantitatively* in that case. However, in one particular case equations (4) or (5) may hold even for such a chain of neurons—while not exactly, yet with good ap-

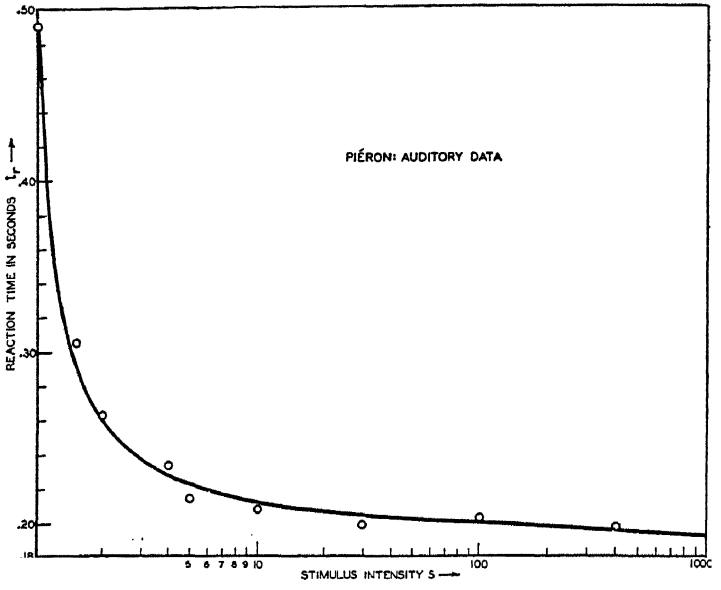


FIG. 33.—Same as Fig. 32 for a set of auditory data by H. Piéron. The values of the constants are: $1/a_e = 0.304$ sec; $H_1 = -0.37$; $H_2 = 0.23$. From H. D. Landahl.[3]

proximation. This will occur when all synapses but one in the chain have, for given conditions, approximately equal and very short delays, t_s, while the one exceptional synapse (or perhaps a very few synapses) has a much larger synaptic delay. Then the total time of transmission along the chain will be controlled mainly by the one "long" synapse, and the relation between stimulus intensity and total transmission time will be of the form given either by equation (4) or by equation (5).

The total time of transmission is then obtained by adding to t_s a constant time, t_0, which represents the duration of the transmission along each fiber, delays at end-organs, etc.

In view of these facts, it may be of interest to compare equations (4) or (5), with a constant term added, to experimen-

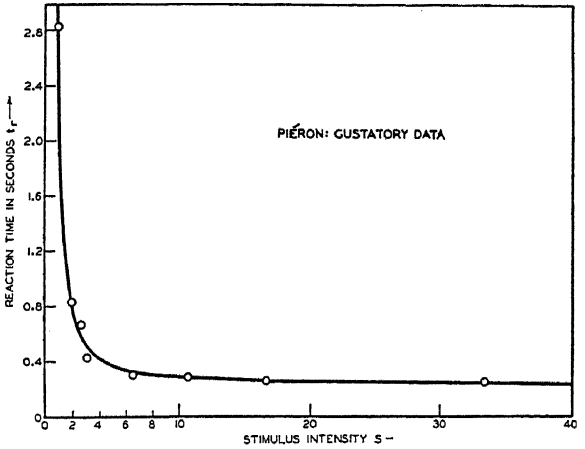


FIG. 34.—Reaction times for gustatory data, plotted against intensity of stimulus. The curve represents the theoretical equation (5). The circles are experimental values by H. Piéron. The values of the parameters in equation (5) are, in this case: $1/a_e = 3.8$ sec; $h_1 = 0.6$; $a_e h_2 / A_e a_1 I_1 = 0.2$. From H. D. Landahl.[3]

tal data for reaction times. Such a comparison has been made by H. D. Landahl,[3] and the results are shown in Figures 32, 33, and 34. Visual and auditory data (Figs. 32 and 33) are well represented by an equation of the form of equation (4), while the simpler equation (5) is sufficient to represent the gustatory data. For discussion of details we must refer to the original paper by H. D. Landahl.[3]

A more complex neuronic structure is represented in Figure 35. Here a neuron *II* of second order is excited by two

neurons, *I* and *III*, of the first order. We may consider cases where the stimuli S_1 and S_3 are applied at different times. Of particular interest is the case where S_3 is too weak to produce a response in neuron *II* and is always applied before S_1. Following H. D. Landahl, we shall consider a special case, namely, where neuron *I* produces only the factor ϵ, while neuron *III* produces both ϵ and j but is of the excitatory type (chap. viii, expressions [7]). Later on we shall discuss a more general assumption regarding neuron *I*. The constants A, a, B, and b are considered as different for the three neurons and are denoted by corresponding subscripts. It is assumed that $A_3/a_3 = B_3/b_3$, although $A_3 > B_3$ and $a_3 > b_3$.

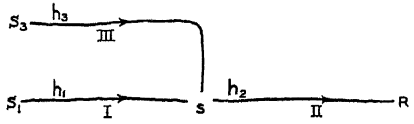


FIG. 35

If a constant stimulus S_1 of indefinite duration is applied suddenly to the neuron *I*, then at the synapse the amount ϵ_1 of ϵ produced by that neuron is given by

$$\epsilon_1 = \frac{A_1 E_1}{a_1}\left(1 - e^{-a_1 t_1}\right), \tag{6}$$

the time t_1 being counted from the moment of arrival of the excitation of *I* at the synapse. Denoting by t_3 the time from the moment of arrival of the excitation of neuron *III* at the synapse, we have, similarly, for the variation of the amounts ϵ_3 and j_3 of the factors ϵ and j produced by neuron *III*:

$$\epsilon_3 = \frac{A_3 E_3}{a_3}\left(1 - e^{-a_3 t_3}\right); \qquad j_3 = \frac{B_3 E_3}{b_3}\left(1 - e^{-b_3 t_3}\right). \tag{7}$$

Excitation of neuron *II* occurs when $\epsilon_1 + \epsilon_3 - j_3 = h_2$. According to our assumption, however, $\epsilon_3 - j_3$ remains always less than h_2. Let us denote by t'_1 the time between the arrival of the excitation of the neuron *I* at the synapse and the initiation of excitation in neuron *II*. In other words, t'_1 is the value of t_1 at the moment when neuron *II* becomes excited. Let t'_3 be the value of t_3 at this moment. (Remember that t_1 and t_3 are counted from different origins, since S_1 and S_3 are not applied simultaneously.) With the foregoing assumption, t'_1 is obtained by introducing expressions (6) and (7) into $\epsilon_1 + \epsilon_3 - j_3 = h_2$, putting $t_1 = t'_1$ and $t_3 = t'_3$, and solving the resulting equation with respect to t'_1. This gives

$$t'_1 = -\frac{1}{a_1} \log \left\{ 1 - \frac{a_1}{A_1 E_1} \left[h_2 - \frac{A_3 E_3}{a_3} (1 - e^{-a_3 t'_3}) + \frac{B_3 E_3}{b_3} (1 - e^{-b_3 t'_3}) \right] \right\}. \quad (8)$$

The total time t_r between the application of stimulus S_1 and the final reaction of the end-organ is obtained under similar assumptions, as before, by adding to t'_1 a constant t_0.

Of particular interest is the study of the relation between the reaction time t_r and the time t'_3 for the case where the intensities of both stimuli are kept constant and where the conduction time along neuron *I* and *III* is very short, as compared with either t'_3 or t'_1. In other words, we consider that most of the constant time t_0 is due to conduction on the efferent side and to delays at the end-organs. In that case, t'_3 approximately represents the time between presentation of the stimulus S_3 and the beginning of the reaction less the time t_0.

These considerations can be applied to some experiments on the effect of a warning or preparatory stimulus upon the length of the reaction time. In such experiments the subject

is to react to a given stimulus; but some time before that stimulus is applied, a different preparatory or warning stimulus is given. We may tentatively identify our stimulus S_1 with the stimulus to which the reaction takes place, while S_3 may be considered as the warning stimulus. In practice the interval t_w between S_3 and S_1 is always much larger than the reaction time t_r. In other words,

$$t_w = t'_3 - t'_1 \gg t_r = t'_1 + t_0 .$$

In that case we have approximately

$$t_w = t'_3 .$$

Putting

$$t'_3 = t_w ; \qquad M = 1 - \frac{a_1 h_2}{A_1 E_1} ; \qquad J = \frac{A_3 E_3 a_1}{A_1 E_1 a_3} = \frac{B_3 E_3 a_1}{A_1 E_1 b_3} , \quad (9)$$

we obtain from expression (8)

$$t_r = t'_1 + t_0 = t_0 - \frac{1}{a_1} \log \left[M + J(e^{-b_3 t_w} - e^{-a_3 t_w}) \right] , \quad (10)$$

which is the relation between the preparatory interval t_w and the reaction time t_r. In Figure 36 are shown two sets of actual data,[3] compared with curves represented by equation (10).

It is of interest to note that the value of $1/a_e$ is of the same order of magnitude for the following curves (Fig. 36) as for the visual data (Fig. 32) and the auditory data (Fig. 33), although there is no apparent connection between the curves. However, it should be noted here that the intensity-time curves are largely determined by the quantity $H_1 + H_2$, the value t_0, and the product of H_2 and $1/a_e$. As long as H_2/a_e remains constant, changes in H_2 or in $1/a_e$ do not appreciably affect the shape of the curve. Then, making the value $1/a_e$

the same for two different situations imposes a condition upon H_2. If, then, any one is determined independently, the

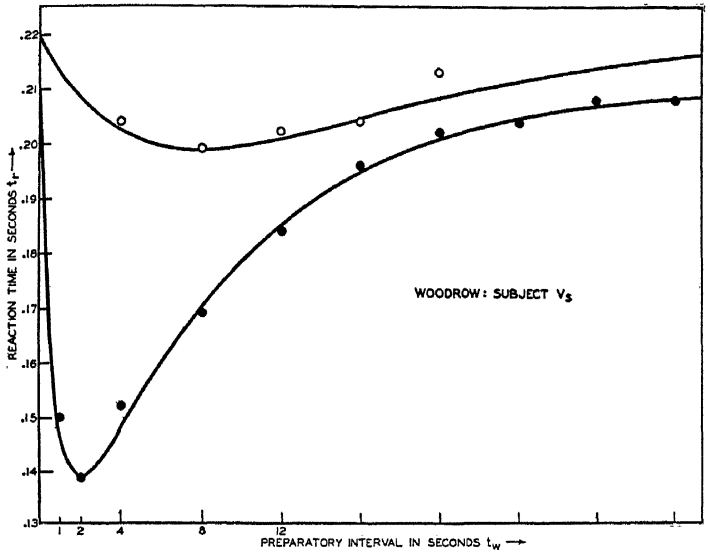


FIG. 36.—Reaction time plotted against the interval between warning and final stimulus. The curves represent equation (10). The points and circles represent observations by H. Woodrow. The two sets of data were obtained on the same subject but under different conditions. For the case represented by circles the subject did not know the lengths of the warning period, whereas in the other cases the subject was given practice with a particular warning period, after which his responses with that warning period were recorded. The values of the parameters are as follows: Upper curve: $t_0 = 0.13$ sec; $1/a_1 = 0.2$ sec; $M = 0.638$; $J = 2$; $a_3 = 0.132$ sec^{-1}; $b_3 = 0.12$ sec^{-1}. Lower curve: $t_0 = 0.13$ sec; $1/a_1 = 0.2$ sec; $M = 0.67$; $J = 0.41$; $a_3 = 1.28$ sec^{-1}; $b_3 = 0.128$ sec^{-1}. From H. D. Landahl.[3]

others are fixed. Thus we see that the curve of equation (4) is, apart from the constant term added, practically a two-parametric curve.

The peculiar minimum of the t_r, t_w curves is due, in the present theory, to the fact that the constants of the neuron *III* satisfy the relations (7) of chapter viii. As has been shown in *MB*, page 227, in this case the difference $\epsilon_3 - j_3$ first increases with t_3', then reaches a maximum, and finally decreases, tending asymptotically to a constant value positive for $A_e/a_e > B_e/b_e$, or zero for $A_e/a_e = B_e/b_e$. Thus, for a certain value of t_3' a maximum amount of $\epsilon_3 - j_3$ is added at the synapse. Since the condition of synaptic transmission

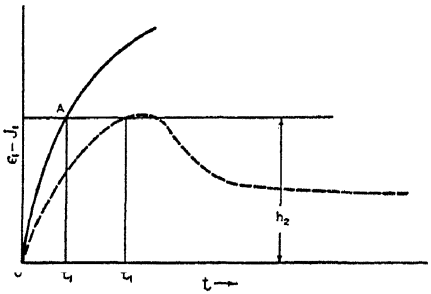


FIG. 37

is $\epsilon_1 + \epsilon_3 - j_3 = h_2$, the larger the $\epsilon_3 - j_3$, the smaller is the necessary amount of ϵ_1 and the sooner will this necessary amount be reached according to equation (6).

If we consider the more general case where neuron *I* also produces both ϵ and j, things become more complicated. However, approximately the same relation will hold for the case where S_1 is sufficiently larger than the threshold necessary to produce any reaction at all. In this general case the difference $\epsilon_1 - j_1$ of ϵ and j produced by neuron *I* will vary with respect to t_1', as shown in Figure 37. When S_1 and, therefore, E_1 are near threshold, then the excitation of neuron *II* begins at the time t_1'' (Fig. 37, dotted line). When S_1 and E_1 are large, then neuron *II* is excited at the time t_1' (Fig. 37, full

line). But in this case the segment OA of the curve may be represented, with sufficient accuracy, by equation (6).

From such simple neuronic structures as are considered here, we may pass to more complex ones. A particularly interesting structure is that consisting of two purely excitatory neurons forming a closed circuit, so that one neuron excites the other, and the latter in its turn excites the first. As has been shown in *MB*, chapter xxiv, and later on in more detail in a paper by A. S. Householder,[4] such a circuit possesses the property of remaining unexcited as long as any external stimuli do not exceed a certain threshold. When, however, an external stimulus exceeds a given threshold, then the circuit is brought irreversibly into a permanently excited state, in which it remains even after the removal of the external stimulus. Applications of such structures to the theory of conditioned reflexes have been made.[5]

A. S. Householder and H. D. Landahl[6] have studied more complex circuits, involving neurons producing both ϵ and j, so that they are not purely excitatory. In that case the permanent excitatory state of the whole system is not necessarily constant, as in the former case. The intensity of excitation may oscillate with respect to time. We thus have a system, based on our fundamental equations, which possesses a property of producing spontaneously, without external stimulation, a rhythmic excitation. Possible applications of such considerations to the theory of the spontaneous rhythmic activity of the brain[7] may be indicated.

It has been pointed out in *MB* (chap. xxii, p. 234) that, when we have a chain of neurons with different thresholds and other constants, the intensity of excitation of the individual neurons of the chain may either increase or decrease along the chain. Landahl[8] has shown that a rather simple relation between thresholds of the individual neurons leads to exponentially increasing or decreasing excitation. House-

holder[9] has studied more complex cases. If the end-link of a chain is an inhibitory neuron, the chain has a net inhibitory effect which may either increase or decrease with the length of the chain.

REFERENCES

1. G. von Bonin, *Psychometrika*, 4, 69, 1939.

2. G. E. Coghill, *Anatomy and the Problem of Behaviour* (Cambridge University Press, 1929).

3. H. D. Landahl, *Bull. Math. Biophysics*, 1, 95, 1939.

4. A. S. Householder, *Psychometrika*, 3, 273, 1938.

5. N. Rashevsky, *Psychometrika*, 2, 199, 1937.

6. A. S. Householder and H. D. Landahl, *Psychometrika*, 4, 255, 1939.

7. Cf. a number of papers on this subject in *Cold Spring Harbor Symposia on Quantitative Biology*, 4, 1936.

8. H. D. Landahl, *Psychometrika*, 3, 291, 1938.

9. A. S. Householder, *Psychometrika*, 3, 69, 1938.

CHAPTER X

DISCRIMINATION OF INTENSITIES

Let a stimulus of intensity S_1 be applied to a sense organ. Under actual physiological conditions, perceptible stimuli, no matter how weak and how sharply localized, excite not one but a large number of peripheral fibers. A nerve innervating a given sense organ usually consists of a large number of fibers with different thresholds. A stimulus of a given intensity, S_1, excites only a fraction of those fibers, namely, those whose thresholds are less than S_1. If we apply to the same sense organ a stronger stimulus of intensity, $S_2 > S_1$, then that second stimulus will excite all the fibers which were excited by the first one, plus an additional number of fibers, namely, those whose thresholds lie between S_1 and S_2. An excessively strong stimulus, such that its intensity exceeds the highest threshold of the bundle of nerve fibers in the nerve trunk, will excite all the fibers.

Whether each peripheral fiber is connected by a chain of neurons with a corresponding single fiber of an effector end-organ, or whether through branches and collaterals this peripheral fiber becomes connected to several fibers of the effector end-organ, thus producing a sort of "multiple response," the foregoing considerations lead to the conclusion that a response due to a weaker stimulus is, so to say, always "contained" in the response for any stronger stimulus of the same type. In this simple scheme, whenever a stimulus, S_1, produces a reaction, R_1, then a stronger stimulus, $S_2 > S_1$, produces necessarily such a reaction, R_2, that it includes the reaction R_1. However, the fact that we can actually discriminate between stimuli of different intensities shows that

the situation is not simple. We respond to a weaker stimulus by calling it "weaker," and to a stronger stimulus by calling it "stronger." The words "weaker" and "stronger" themselves constitute certain motor reactions of the lungs, pharynx, and tongue to the corresponding stimuli; and neither of these reactions "includes" the other in the above-mentioned sense. A still better illustration is obtained by considering cases where we are taught to respond to a weaker stimulus in a way qualitatively different from our response to a stronger one. Thus, a weak sound of a given pitch may be used as a

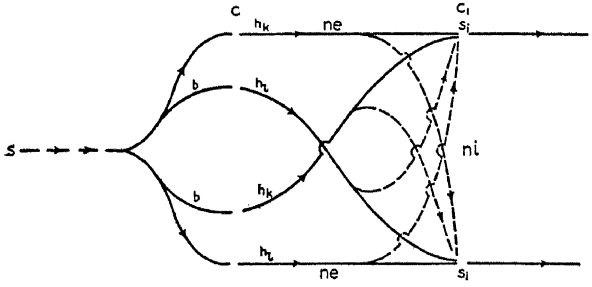


FIG. 38

signal for beginning to read a book, while a stronger sound of the same pitch may be used as a signal for an entirely different action.

Hence, some neurological mechanism must exist which provides for the possibility of a *qualitatively* different response to *quantitatively* different stimuli. Such a mechanism can be provided by the existence of inhibitory fibers, intercalated in a proper way between excitatory ones.

Let us consider the following structure (Fig. 38) and let us confine ourselves to stationary states—that is, we shall consider constant stimuli of sufficiently long duration so that at all synapses the factors ϵ and j have practically reached their constant limiting values. In that case ϵ is proportional to E

(cf. Eq. [1] of chap. ix). A sensory peripheral fiber or a chain of fibers eventually divides into several branches. Each of these branches synapses with a neuron *ne* of higher order. The thresholds h of the neurons *ne* are, in general, different and are distributed according to some distribution function, $N(h)$. That is, if we take h as abscissa and plot the number $N(h)$ of neurons *ne* having a threshold between h and $h + dh$, where dh is a very small fixed quantity, we shall obtain a curve of some sort, which represents $N(h)$ as a function of h. If, as in *MB*, chapter xxii, we consider the simpler case where, whenever a fiber divides into several branches, the intensity E of excitation in each branch is the same as in the original fiber, then we have the following situation.

For a given intensity S of the peripheral stimulus there corresponds a definite value E_1 of intensity of excitation in each branch (*MB*, chap. xxii). If the chain consists of only one fiber, then E_1 is a linear function of S. In general, however, E_1 is a more complicated function of S, depending on the constants of the individual members of the chain

$$E_1 = F(S) . \tag{1}$$

The intensity of excitation E_{ne} of each neuron *ne* is given by (*MB*, chap. xxii)

$$E_{ne} = \beta(PE_1 - h) . \tag{2}$$

The total intensity of excitation of all neurons (*ne*) which have a threshold $h < PE_1$ is then given by

$$E(h) = \beta(PE_1 - h)N(h) . \tag{3}$$

If $N(0) = 0$, as should be expected from general considerations, then $E(h)$ is zero for $h = 0$ and for $h = PE_1$ and is positive within this interval. Hence, in that interval $E(h)$ has

at least one maximum. Let us first consider the case where it has only one maximum, $h = h_m$, h_m being a function of E_1.

Let all neurons *ne* send fibers to a region of the brain, C_1, in which all neurons *ne* with the same threshold h form a synapse with the same neuron of third order ne_3; and let, moreover, each neuron *ne* of threshold h_k send off a collateral which excites inhibitory fibers *ni*, leading to all the synapses s_i corresponding to neurons of a different threshold h_l. Consider, for simplicity, that all thresholds h_i of *ni* are the same. Each synapse s_i is excited by an amount of $(\epsilon - j)_1$ due to neurons *ne* and proportional to the quantity $E(h)$ in equation (3) (cf. chap. ix, p. 135). Moreover, each of these synapses receives from neurons *ni* a certain amount of $(j - \epsilon)_2 > 0$. Since each neuron *ni* is excited by a neuron *ne*, its intensity of excitation is stronger, the stronger the intensity of excitation of the corresponding neuron *ne*. The few synapses, s_i, corresponding to such a value of h that $E(h)$ is very high, will receive a large amount of $(\epsilon - j)_1$ from the *ne* neurons and relatively lesser amounts $(j - \epsilon)_2$ from the inhibitory fibers, coming from other less excited neurons *ne*. The total amount of $\epsilon - j$ will therefore be positive and large enough, and the corresponding neurons ne_3 will be excited. But the large number of synapses, corresponding to such values of h for which $E(h)$ is much smaller will receive little $(\epsilon - j)_1$ from their *ne* neurons and a large amount of $(j - \epsilon)_2$ from the strongly excited neurons *ne*. As a result of this, only those synapses which correspond to sufficiently large values of $E(h)$ will transmit excitation to neurons ne_3. But such $E(h)$ has a maximum for a value h_m of h_1, and therefore only those synapses s_i will transmit excitation that correspond to values of h, that lie in the neighborhood of h_m, and that are therefore included between two fixed values h_1 and h_2 (Fig. 39). Since, according to equation (3), h_m is a function of E_1 and hence also a function of S, the h_1 and h_2 are also func-

tions of S. If, by varying the intensity S of the peripheral stimulus, we vary E_1, this will result in a variation of h_m; and if, for a new value E_1' of E_1, the corresponding h_m' will be sufficiently different from the h_m, then entirely different groups of synapses s_i will be excited by the stimulus S' than by S. Thus, to any intensity S of the peripheral stimulus there is a corresponding excitation of a definite group of synapses s_i. A stimulus of intensity S will produce a reaction R through a group s_i of synapses, while a stimulus of a different intensity, S', will not produce R because it involves a totally different group, s_i' of synapses. Each intensity S of the same stimulus has thus a representative individual group

FIG. 39

of synapses in the nerve centers, and therefore *each intensity may, in a way, be considered as a different stimulus-pattern.*

However, things will happen in this fashion only if the difference between the intensities S and S' is sufficiently large. If this is not the case, then h_m and h_m' differ very little, and the corresponding intervals (h_1, h_2) and (h_1', h_2') will partially overlap (Fig. 39). If, now, S produces reaction R, then S' also produces R through all the synapses which correspond to the interval (h_1', h_2) (Fig. 39). Since, because of equation (3), h_1 and h_2, as well as h_m, are functions of E_1, or, what amounts to the same, functions of S, the minimum difference $\Delta S = S - S'$ for which the intervals (h_1, h_2) and (h_1', h_2') do not overlap at all is itself a function of S. The determination of this function $\Delta S = U(S)$ is, in principle, a simple problem, which may, however, involve some complicated algebra. The function $U(S)$ is determined by $F(S)$ in equation (1) and by $N(h)$. Several problems suggest them-

selves at this stage. For instance, we may ask the following question. How should $F(S)$ and $N(h)$ be chosen in order that $\Delta S = U(S)$ would have a prescribed form—say that of Fechner's Law?

When (h_1, h_2) and (h_1', h_2') partially overlap, we may investigate the relative intensity of R, as produced by S', compared to that produced by S. In the simplest case this relative intensity will be given by the ratio η of all neurons lying in the interval (h_1', h_2) (Fig. 39) to those lying in the interval (h_1, h_2). That is,

$$\frac{R'}{R} = \eta = \frac{\int_{h_1'}^{h_2} N(h)dh}{\int_{h_1}^{h_2} N(h)dh}. \tag{4}$$

If we make more complicated assumptions about the possible interaction of the neurons which lie on the efferent side of the synapses s_i, we shall obtain expressions different from (4). This leads to another interesting group of problems.

The case where $E(h)$ has several maxima in the interval $(0, PE_1)$ is more complicated but is treated in a similar way.

A similar but slightly different situation is obtained by considering the case where the inhibitory fibers are not collaterals of the *ne* neurons but are themselves excited at the synapses s_i (Fig. 40). This case has been discussed in more detail in *MB*, chapter xxii, and has been used by A. S. Householder to develop a quantitative theory of discrimination along the lines outlined above. We shall now follow Householder's presentation.[1]

The quantity $E(h)$ in equation (3) is a function of h and of the parameter S. Let us denote it by $\Phi(S, h)$. Assuming that the relation (1) is a linear one, we can bring equation (3), by a suitable choice of units, into the form

$$\Phi(S, h) = (S - h)f(h). \tag{5}$$

Since ϵ is proportional to E, we may, by a proper choice of units, make $\epsilon = \Phi$. Let each synapse $s_{h'}$ be connected with each other synapse s_h by an inhibitory fiber. Then the *net amount* $\epsilon - j$ at s_h, to be denoted by $\sigma(S, h)$, will be obtained by subtracting from $\Phi(S, h)$ the amount of j produced by all inhibitory fibers leading from other synapses to s_h. Moreover, $\sigma(S, h)$ is the effective stimulus acting upon all inhibitory fibers leading from s_h. To set up the equation giving σ, we make the following *assumptions concerning the inhibitory*

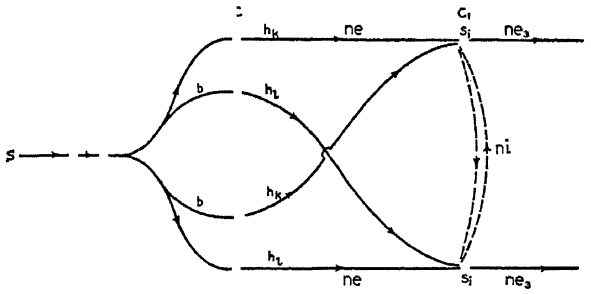


FIG. 40

fibers: (*a*) they are all similar; (*b*) for each fiber the amount of j produced is a linear function of σ; and (*c*) the threshold of each fiber is negligible.

Then the total amount of j *produced by these fibers at any synapse* s_h *is*

$$I(S) = \lambda \int \sigma(S, h) dh ,$$

where λ is a constant of proportionality measuring the activity of the inhibitory fibers. The integration is to be extended over all values of h for which $\sigma > 0$. Hence,

$$\sigma(S, h) = \Phi(S, h) - \lambda \int \sigma(S, h) dh .$$

We may suppose $f(h)$ to be at least continuous and to vanish at $h = 0$. Then $\Phi(S, h)$ vanishes at $h = 0$ and at

$h = S$, for any given S. Hence, $\Phi(S, h)$ has at least one maximum in the interval. If we suppose it to have only one maximum, then for any S the graph of Φ will have just two points of ordinate I—say at h_1 and at h_2—and for values of h between these two values σ will be positive. Call this interval over which $\sigma > 0$ the "excited interval."

Putting

$$I(S) = \lambda \int_{h_1}^{h_2} \sigma(S, h) dh \,, \tag{6}$$

we find

$$\sigma(S, h) = \Phi(S, h) - I(S) \,;$$

and hence,[2]

$$I(S) = \frac{\lambda \int_{h_1}^{h_2} \Phi(S, h) dh}{1 + \lambda(h_2 - h_1)} \,. \tag{7}$$

For determining $h_1(S)$ and $h_2(S)$ we have, then,

$$\Phi(S, h_1) = \Phi(S, h_2) = \frac{\lambda \int_{h_1}^{h_2} \Phi(S,h)\, dh}{1 + \lambda(h_2 - h_1)} \,, \tag{8}$$

since at these points σ vanishes.

We now define the Weber ratio $\delta(S)$ by the equation

$$h_2(S) = h_1(S + S\delta) \,. \tag{9}$$

This is equivalent to saying that discrimination between the intensities S and $S(1 + \delta)$ is possible (in a suitable percentage of trials) when the excited intervals are just distinct.

By simple considerations A. S. Householder proves[1] that for very large values of S equation (9) is inconsistent with the foregoing assumption. In other words, the above-described

mechanism does not work when S is excessively large. The limits of possible variations of S are approximately defined by the value h^* corresponding to the maximum of $f(h)$. Neurophysiologically, there is nothing unlikely in that result, inasmuch as any observed regularities hold with sufficient exactness only within more or less limited ranges of stimulus intensities.

Under those conditions we need consider only the ascending branch of $f(h)$. Since nothing is known about that function, we may make the simplest assumption, namely, that over a wide range of values of S, the function $f(h)$ is, with sufficient approximation, linear, so that

$$\Phi(S, h) = h(S - h) . \tag{10}$$

In that case the graph of $\Phi(S, h)$ is an inverted parabola with a maximum at $S/2$. Hence, h_1 and h_2 are equidistant from $S/2$. Define the "relative interval," x, by

$$Sx = S - 2h_1 = 2h_2 - S . \tag{11}$$

Thus,

$$\Phi(S, h_1) = \Phi(S, h_2) = \frac{S^2(1 - x^2)}{4} \tag{12}$$

$$\int_{h_1}^{h_2} \Phi(S, h)dh = \frac{S^3x(3 - x^2)}{12} \tag{13}$$

Then equation (8) becomes

$$2\lambda Sx^3 + 3x^2 - 3 = 0 . \tag{14}$$

If we set

$$u = 2\lambda S \tag{15}$$

we obtain

$$ux^3 + x^2 - 1 = 0 . \tag{16}$$

To calculate δ we use equation (9). Let $\bar{x}$ and $\bar{u}$ be the values of x and u corresponding to the value $S(1 + \delta)$ of the stimulus. Then we have

$$\bar{u} = u(1 + \delta) . \tag{17}$$

From equations (9) and (11) and the definition of $\bar{x}$ we obtain, after simple transformations,

$$\bar{x} = \frac{\delta - x}{\delta + 1} . \tag{18}$$

Then by defining a new variable z by the equation

$$\delta = x + z , \tag{19}$$

we obtain

$$uz^3 - 2(x + 1)z - (x + 1)^2 = 0 . \tag{20}$$

Eliminating z from equations (19) and (20), we shall find an equation connecting u, δ, and x. Then eliminating x from that equation and from equation (16), we obtain a relation between δ and u. Since u is connected to the intensity S of the stimulus by means of equation (15), being actually proportional to S, we thus arrive at a relation between the Weber ratio δ and the stimulus intensity S.

For further details and discussion of the foregoing equations we must refer the reader to the original paper by A. S. Householder.[1] Figures 41–45, taken from Householder's paper, show a comparison of the calculated and observed relations between δ and S (or, what amounts to the same, between δ and u). For convenience δ is plotted not against u

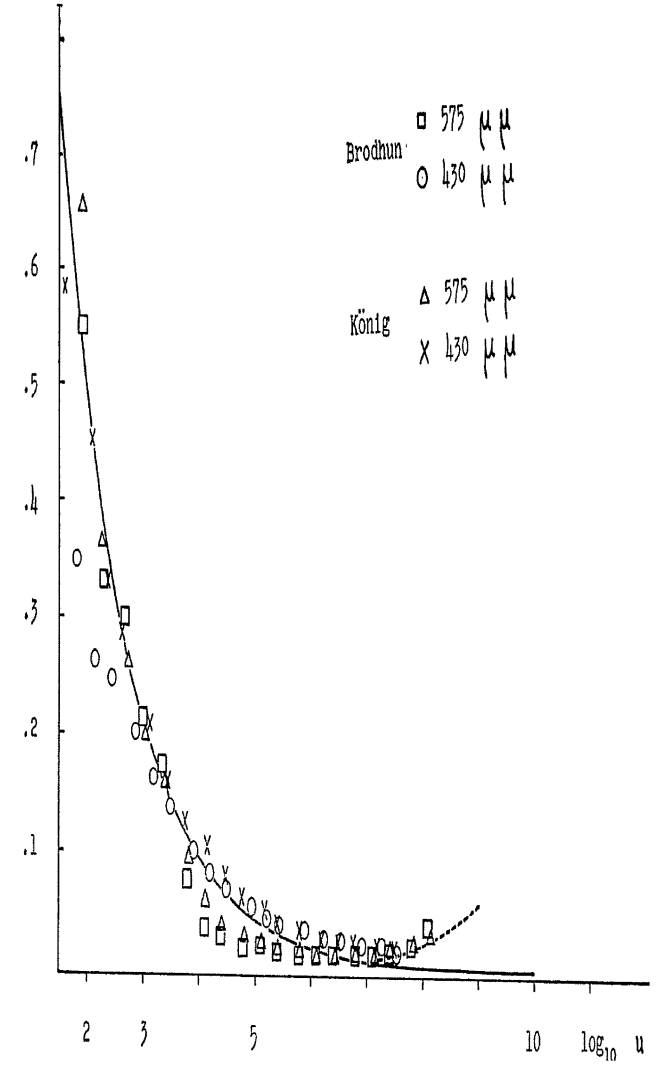


FIG. 42.—Same as Fig. 41 but for another set of data

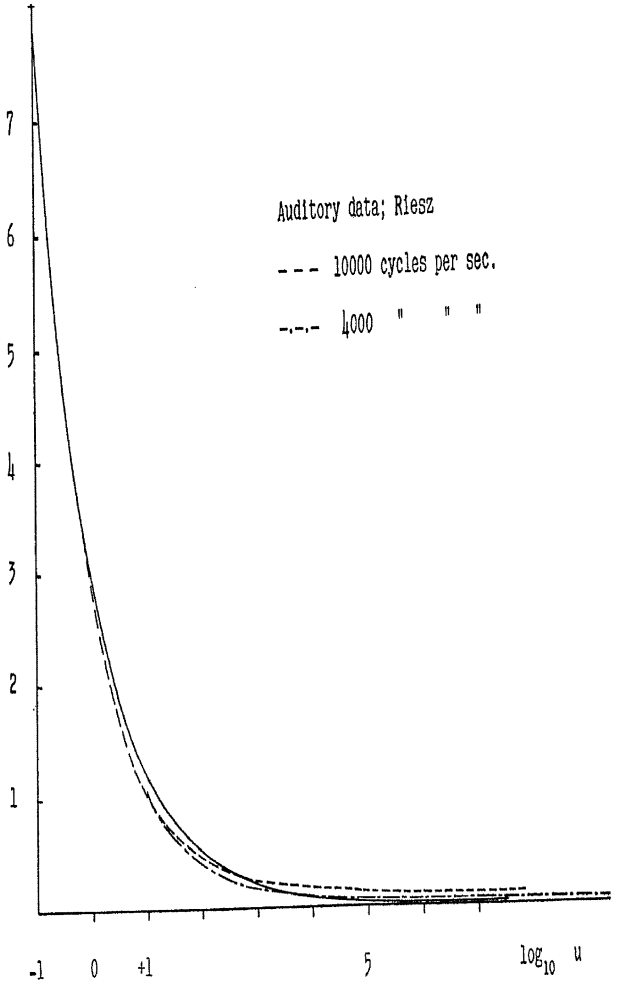


FIG. 43.—The full line represents the theoretical relation between the Weber ratio δ, as defined by equation (9), and the quantity u, as defined by equation (15). The broken and alternate lines represent relations found experimentally by R. Riesz for auditory stimuli. From A. S. Householder.[1]

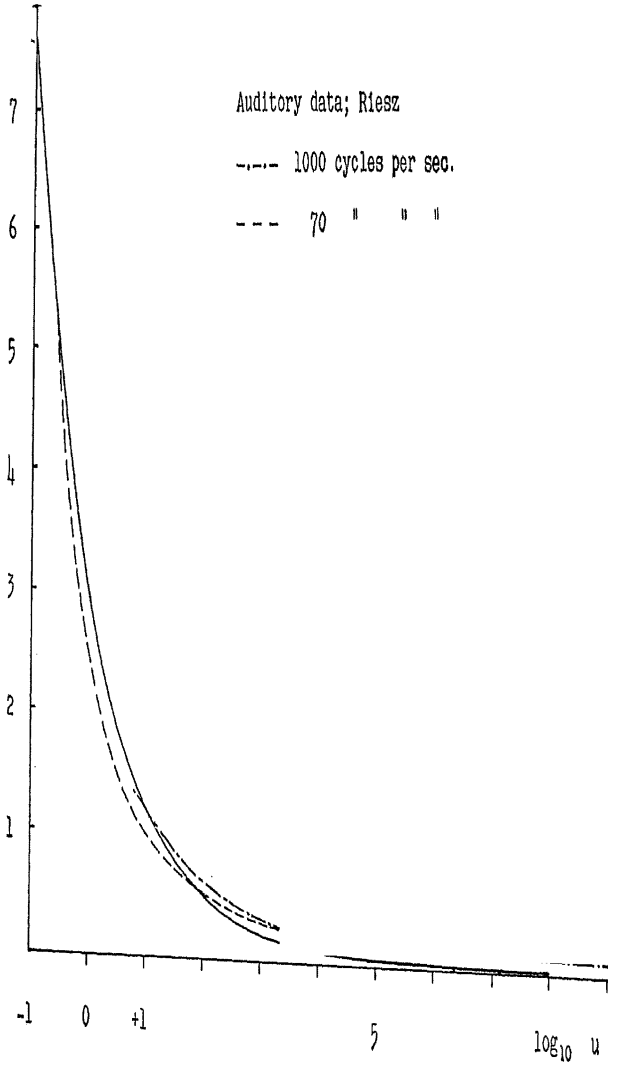


FIG. 44.—Same as Fig. 43 but for another set of data

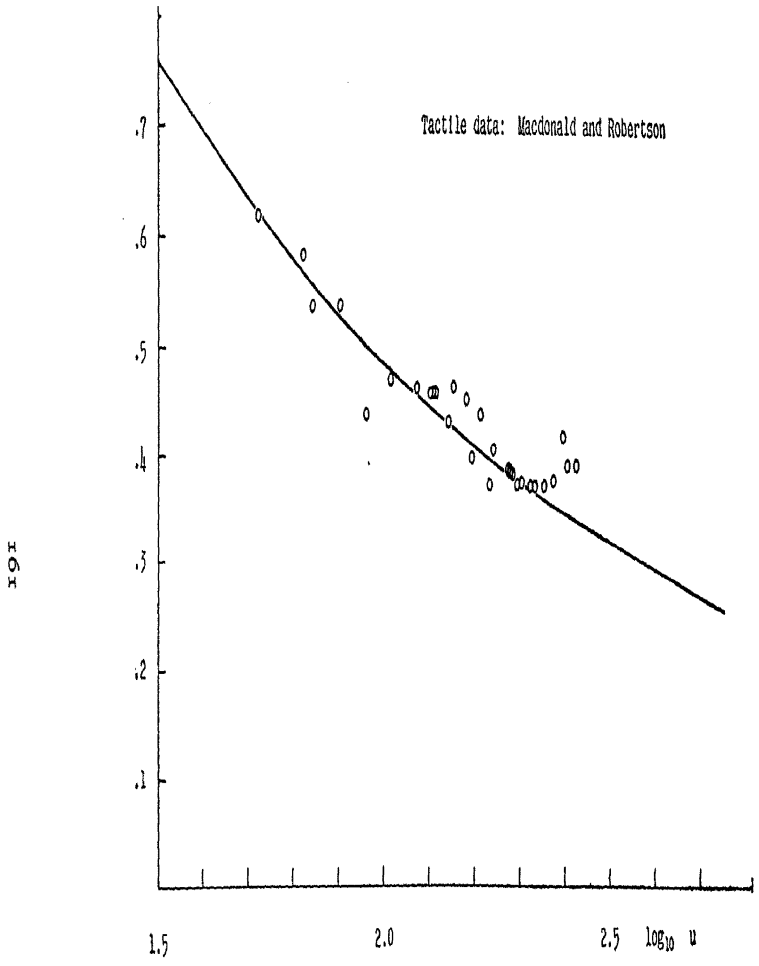


FIG. 45.—The curve represents the theoretical relation between the Weber ratio δ, as defined by equation (9), and the quantity u, as defined by equation (15). The circles represent experimental data for tactile stimuli. From A. S. Householder.[1]

but against $\log_{10} u$. From equations (15), (16), and (20) it follows that only one parameter, namely, the quantity λ, is involved in the final relation between δ and S.

If the relation (1) between E_1 and S is not a linear one, then equation (5) is not strictly equivalent to equation (3). If $F(S)$ in equation (1) is of the type discussed in chapter viii, and if the deviation from linearity is small, within the range of values of S used, then a proper correction applied to equation (5) will result, as shown by Householder,[1] in a slight upturn of the δ, $\log_{10} u$ curve, as indicated on Figures 41 and 42 by the broken line. As will be noticed, the experimental points also indicate such an upturn.

REFERENCES

1. A. S. Householder, *Psychometrika*, 4, 45, 1939.

2. A. S. Householder and E. Amelotti, *Psychometrika*, 2, 255, 1937.

CHAPTER XI

MATHEMATICAL BIOPHYSICS OF PSYCHOPHYSICAL DISCRIMINATION

Hitherto we have discussed the problem that arises when we ask whether two intensities which are only slightly different from each other are definitely perceived as different or not. This problem, however, has another aspect. The two intensities, or, more generally, the two stimuli, may be so close to each other that a definite discrimination is impossible. Yet, even then, there may be a certain probability for a correct statement with regard to whether the two stimuli are equal or not. The problem acquires a still slightly different aspect when we ask not merely for a statement whether the stimuli are equal or different but also for a statement as to which of the two stimuli is the greater (or the stronger) one. The closer the two stimuli are to each other with respect to their intensities the less will be the probability for a correct judgment. We may ask for the functional relation between the difference of the two stimuli and the probability of a correct judgment.

An interesting approach to this problem has been made by H. D. Landahl[1] by considering the neurological structure shown in Figure 46. This structure is a particular instance of the structure considered in *MB*, chapter xxii.

Let the stimulus S_1 (Fig. 46) elicit a reaction R_1 through a chain of excitatory fibers connected by synapses s_1 and s_3, while the stimulus S_2 elicits the reaction R_2 through synapses s_2 and s_4. However, let the synapses s_1 and s_2 excite also inhibitory fibers which are arranged as shown on Figure 46.

We have essentially the well-known scheme of reciprocal innervation.

For simplicity, let us consider the case where the excitatory fibers are of purely excitatory type (chap. viii, p. 124), while the inhibitory ones are of a purely inhibitory type.

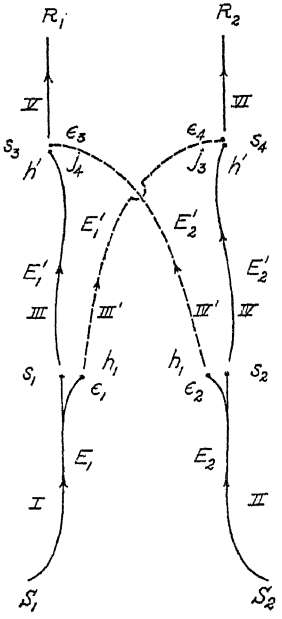


FIG. 46

That is, the former produce only the ϵ-factor; the latter, only the j-factor. This restriction is not an essential one.

The stimuli S_1 and S_2 result in the production, at the synapse s_3, of an amount ϵ_3 of ϵ by fiber *III* and an amount j_4 of j by the fiber *IV'*. Similarly, at the synapse s_4 we have the amounts ϵ_4 and j_3 when the stimuli S_1 and S_2 are applied. If the constants of all the fibers are the same, and in particular when $A = B$ and $a = b$, then we always have $\epsilon_3 = j_3$ and

$\epsilon_4 = j_4$. When $S_1 = S_2$, then at the synapses s_3 and s_4 we have, correspondingly,

$$\epsilon_3 = \epsilon_4 ; \qquad j_3 = j_4 ;$$

and hence,

$$\epsilon_3 - j_4 = \epsilon_4 - j_3 = 0 . \tag{1}$$

In other words, when the two stimuli S_1 and S_2 are equal, they mutually inhibit each other, and neither R_1 nor R_2 is produced.

If, however, one of the stimuli—for instance, S_1—is much larger than the other, S_2, then the amount ϵ_3 is increased, resulting in an increase of $\epsilon_3 - j_4$. At the same time, j_3 is also increased, resulting in a decrease of $\epsilon_4 - j_3$. Therefore, when $\epsilon_3 - j_4$ becomes so large as to exceed the threshold h' of fiber V, R_1 will be elicited, but R_2 will still be inhibited because $\epsilon_4 - j_3 < 0$. For a perfectly symmetrical scheme, like the one shown in Figure 46, we have, quite generally,

$$\epsilon_4 - j_3 = - (\epsilon_3 - j_4) . \tag{2}$$

Therefore, at only one of the synapses s_3 or s_4 can $\epsilon - j$ be positive. Hence, the simultaneous presentation of S_1 and S_2, regardless of their intensities, results either in no reaction at all ($S_1 = S_2$) or only in one reaction (R_1 if $S_1 \gg S_2$; R_2 if $S_1 \ll S_2$).

Instead of speaking of the stimuli S_1 and S_2, we may speak of the amounts ϵ_1 and ϵ_2 of ϵ at the synapses s_1 and s_2, since these amounts are monotonically increasing functions of S_1 and S_2, respectively. In order to produce the response R_1, the quantity $\epsilon_3 - j_4$ at the synapse s_3 must exceed the threshold h'. But in order that this should happen, we must have a sufficient excess of ϵ_1 over ϵ_2; in other words, $\epsilon_1 - \epsilon_2$ must exceed a threshold value h, so that

$$\epsilon_1 - \epsilon_2 > h . \tag{3}$$

The quantity h is not to be confused with the threshold h' of the fibers *III*, *IV*, *V*, and *VI* or with the threshold h_0 of the fibers *I* and *II*. The quantity h is a function of h', but it involves other parameters also.

This situation holds only as long as all neurophysiological processes are perfectly regular and constant, not being subject to any external or internal disturbances. However, as we have discussed in *MB*, chapter xxviii, in general we must expect spontaneous fluctuation of excitation to occur in the central nervous system. Such fluctuations may be due to fluctuations of metabolic activity or to excitation carried to a given region from a number of other regions of the brain, which are randomly excited by the stream of oncoming exteroceptive as well as proprioceptive and enteroceptive stimuli.

Let us consider such fluctuations at the synapses s_1 and s_2. They result in an addition of a varying amount of $\epsilon - j$ to either ϵ_1 or ϵ_2. In the absence of fluctuations, whenever we have $\epsilon_1 - \epsilon_2 > h$, the reaction R_1 is produced. Since an increase of ϵ_1 is equivalent to a reduction of ϵ_2, as far as the release of R_1 is concerned, we may consider fluctuations of ϵ_1 at the synapse s_1 only.

In the presence of such fluctuations the following will generally occur. Suppose $S_1 \gg S_2$; then $\epsilon_1 \gg \epsilon_2$. In the absence of fluctuations this will result in the reaction R_1. But if, owing to the fluctuations, an amount $(\epsilon - j)' < 0$ is added to ϵ_1 at the synapse s_1, then if the absolute value of this amount $(\epsilon - j)'$ is sufficiently large, the net $\epsilon - j$ at the synapse s_1 will become less than h; and, therefore, either no reaction at all will be produced or a reaction R_2 will occur because of a decrease of j_3 at the synapse s_4, in spite of the fact that $S_1 > S_2$. Let us use the designation "correct response" for the response R_1 when $S_1 > S_2$, or for R_2 when $S_1 < S_2$; and the term "wrong response" for the reaction R_1 when $S_1 < S_2$ or for R_2 when

$S_1 > S_2$. With this terminology we see that the spontaneous fluctuations of excitation may result in wrong responses. The probability of a large fluctuation $(\epsilon - j)'$ is smaller than the probability of a small fluctuation. Since, when $S_1 > S_2$, the larger the $S_1 - S_2$ the larger must be the absolute value of the additional $(\epsilon - j)' < 0$ in order to produce the wrong response; therefore the probability of a wrong response increases when the difference $S_1 - S_2$ decreases.

Quantitatively we can determine the relation between the probability of a given response and the difference of the stimuli in the following way.[1,2] From the preceding discussion it follows that, in order to produce a reaction R_1, we must have (cf. Eq. [3])

$$[\epsilon_1 + (\epsilon - j)'] - \epsilon_2 > h\,. \tag{4}$$

The wrong reaction is produced when the left-hand side of (4) is less than $-h$. When the left-hand side of (4) lies between $-h$ and $+h$, no response is produced.

Hence,

I. If $(\epsilon - j)' > -(\epsilon_1 - \epsilon_2 - h)$, the correct response is made;

II. If $(\epsilon - j)' < -(\epsilon_1 - \epsilon_2 + h)$, the wrong response is made; and

III. If $-(\epsilon_1 - \epsilon_2 - h) \geqq (\epsilon - j)' \geqq -(\epsilon_1 - \epsilon_2 + h)$, no response is made at all. In other words, things occur as if $S_1 = S_2$. We shall call this last case "equality response."

Let

$$p(\epsilon - j)'\, d(\epsilon - j)'$$

denote the probability of having the fluctuation lie between $(\epsilon - j)'$ and $(\epsilon - j)' + d(\epsilon - j)'$. We have

$$\int_{-\infty}^{+\infty} p(\epsilon - j)' d(\epsilon - j)' = 1\,. \tag{5}$$

Since the correct response is made whenever case I is satisfied, the probability P_c of the correct response is equal to the probability of having $(\epsilon - j)'$ lie in the interval $[-(\epsilon_1 - \epsilon_2 - h), +\infty]$, a probability given by

$$P_c = \int_{-(\epsilon_1-\epsilon_2-h)}^{\infty} p(x)dx\,; \qquad x = (\epsilon - j)' \tag{6}$$

Similarly, we have for the probability P_w of the wrong response and for the probability P_e of the equality response from cases II and III, respectively,

$$P_w = \int_{-\infty}^{-(\epsilon_1-\epsilon_2+h)} p(x)dx \tag{7}$$

and

$$P_e = \int_{-(\epsilon_1-\epsilon_2+h)}^{-(\epsilon_1-\epsilon_2-h)} p(x)dx\,. \tag{8}$$

If the function $p(x)$ is given, then equations (6), (7), and (8) give us P_c, P_w, and P_e in terms of ϵ_1, ϵ_2, h, and of any parameters of the function $p(x)$. On the other hand, the quantities ϵ_1 and ϵ_2 can be expressed in terms of intensities of constant stimuli S_1 and S_2 and the duration, t, of their application. In fact, denoting the intensities of excitation of fibers *I* and *II* (Fig. 46) by E_1 and E_2, respectively, we have (chap. viii)

$$\begin{aligned} \epsilon_1 &= \frac{AE_1}{a}(1 - e^{-at})\,; \\ \epsilon_2 &= \frac{AE_2}{a}(1 - e^{-at})\,. \end{aligned} \tag{9}$$

Using now, together with equations (9), the relations (2) and (3) of chapter viii, which gives, approximately,

$$\begin{aligned} E_1 &= K \log \frac{S_1}{h_0}\,, \\ E_2 &= K \log \frac{S_2}{h_0}\,, \end{aligned} \qquad K = Iah_0 \tag{10}$$

where h_0 is the threshold of fibers *I* and *II*, we find relations between ϵ_1 and S_1 and ϵ_2 and S_2. Thus, equations (6), (7), and (8) express the probabilities P_c, P_w, and P_e, in terms of S_1, S_2, the time t and the different parameters entering into the function $p(x)$. In the experiment to be described later, to which some of the present considerations are applied, the time t is kept constant.

If we assume for $p(x)$ the normal distribution function[3]

$$p(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-x^2/2\sigma^2}, \tag{11}$$

then equations (6), (7), and (8) express P_c, P_w, and P_e in terms of S_1, S_2, and the two parameters h and σ. In this case, if we take two observed values of either P_c, P_w, or P_e for any two pairs of S_1 and S_2, we can calculate h and σ and then calculate the values of P_c, P_w, and P_e for any other pair S_1, S_2. Since in the case of a normal distribution (11) the integrations in equations (6), (7), and (8) cannot be made in closed form, numerical tables of the probability integral have to be used.

Other distribution functions, such as, for instance,

$$p(x) = \frac{\sigma_1}{2} e^{-\sigma_1|x|} \tag{12}$$

may also be considered. In equation (12) σ_1 is a constant and $|x|$ denotes the absolute value of x. In this case the integrals in equations (6), (7), and (8) can be calculated in closed form. A comparison of the functions (11) and (12) is shown in Figure 47.

These theoretical results have been applied by H. D. Landahl[1] to the cases of discrimination of weights. Table 1 shows the experimental data. This table means, for instance, that when a standard weight of 200 gm (S_1) is compared with another one of 195 gm (S_2), in 15 per cent of the judgments by a subject the actually greater weight was estimated as

being smaller, in 25 per cent both weights were estimated as equal, and in 60 per cent the correct estimate was made. The second lines, corresponding to the weights in the first column, are obtained by modifying the experiment somewhat, namely, by allowing only "greater-than" or "less-than" judgments, excluding the equality judgments. In this case, for instance, when a comparison of the 195-gm weight with the standard

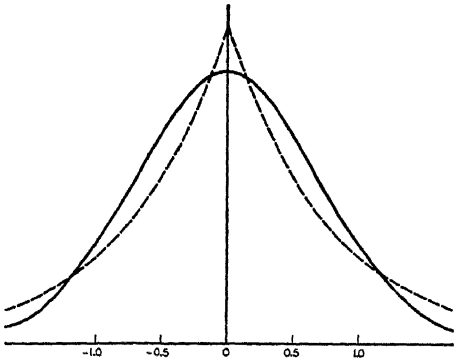


FIG. 47.—The full line represents the normal distribution function, equation (11). The broken line represents the function given by equation (12). The areas of the two curves on this figure are equal.

200-gm weight is made, it is found that there are 20 per cent wrong judgments and 80 per cent correct ones.

The theoretical argument leading to this second mode of experimenting needs a little more discussion. Whenever condition III is satisfied, neither R_1 nor R_2 are produced—in other words, the subject does not state either stimulus to be greater than the other. If, however, the subject is instructed not to make any equality judgments but always to decide on a "greater-than" or "less-than" judgment, this will result in an effort to sharpen his own acuity for the case, when he otherwise would pronounce an equality judgment. Such an effort

may be interpreted biophysically as a general increase in excitability, resulting in a lowering of the threshold h. When $h = 0$, then from equation (8) we have $P_e = 0$. For $h \gtrless 0$, we find from equations (5), (6), (7), and (8)

$$P_c + P_w + P_e = \int_{-\infty}^{+\infty} p(x)dx = 1 . \qquad (13)$$

For $h = 0$, we have

$$P_c + P_w = 1 , \qquad (14)$$

as we should expect from physical considerations.

TABLE 1

EXPERIMENTAL DATA

S (Grams)	P_w	P_e	
185.....	0.05	0.04	0.91
	.05		(.95)
190.....	.12	.18	.70
	.15		.85
195.....	.15	.25	.60
	.20		.80
200.....	.30	(.42)	.28
	.52		.48
205.....	.55	.35	.10
	.85		.15
210.....	.70	.18	.12
	.85		.15
215.....	.85	0.09	.06
	0.93		0.07

TABLE 2

THEORETICAL VALUES

S (Grams)	P_w	P_e	P_c
185.....	0.03	0.06	0.91
	.05		.95
190.....	.07	.13	.80
	.11		.89
195.....	.14	.27	.59
	.24		.76
200.....	.29	.42	.29
	.50		.50
205.....	.58	.28	.14
	.76		.24
210.....	.79	.14	.07
	.88		.12
215.....	.89	0.07	.04
	0.94		0.06

In using data for "two-category" judgments only, we put $h = 0$ in our equations. In this way we can calculate σ from any one pair S_1, S_2.

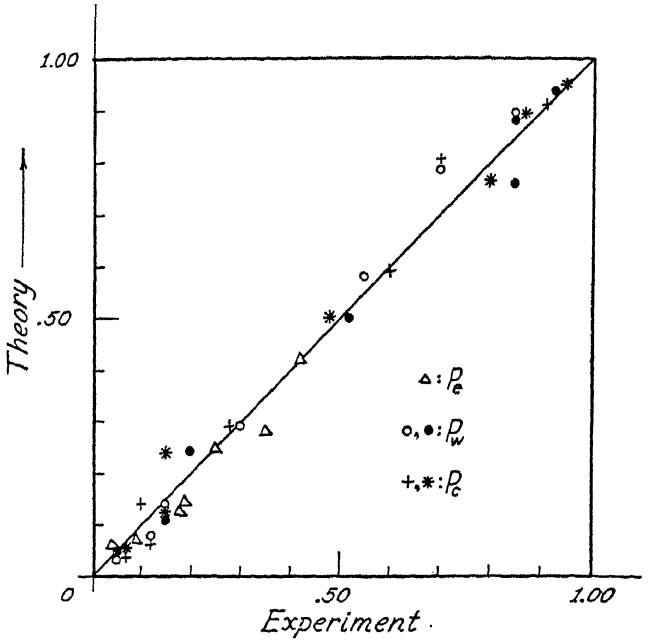


FIG. 48.—Graphs showing the comparison between the experimental data in Table 1 and the theoretical data in Table 2, for probabilities of correct, wrong, and equality judgments. A perfect agreement would bring all points onto the straight line. From H. D. Landahl.[1]

The values in parentheses in Table 1 have thus been used for computing h and σ_1, assuming for $p(x)$ the distribution function (12). With the values thus found, all the remaining probabilities for other values S_1, S_2 were calculated; Table 2 shows these calculated values. A comparison of the two tables is best made graphically and is shown in Figure 48, in

which the theoretically calculated values are plotted against the experimental values.

From equations (6), (7), and (8) it is seen that P_c, P_w, and P_e are functions of the difference $\epsilon_1 - \epsilon_2$. For a constant value t of the duration of stimuli, such as is used experimentally, we see from equation (9) and (10) that therefore P_c, P_w, and P_e are functions of the ratio S_1/S_2.

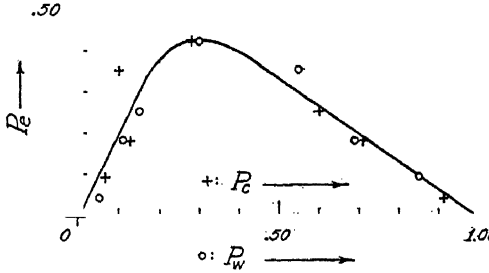


FIG. 49.—The curve represents the theoretical relation between the probability of a correct or wrong judgment and that of an equality judgment (cf. Eq. [14]). The circles and crosses represent experimental data by J. P. Guilford.[4] From H. D. Landahl.[1]

Hence, for $p(x)$, given by equation (12),

$$P_c = f_1\left(\frac{S_1}{S_2}, h, \sigma_1\right); \qquad P_e = f_2\left(\frac{S_1}{S_2}, h, \sigma_1\right). \qquad (15)$$

Since for a given experimental setup h and σ_1 are constants, therefore elimination of S_1/S_2 from the two equations (14) gives us a relation between P_c and P_e. Similarly a relation is obtained between P_w and P_e. The theoretical relation obtained by numerical evaluation of the integrals involved is compared with experimental data in Figure 49.

For a different set of data on weight discrimination, similar comparison was made by using a normal distribution func-

tion, represented by equation (11). The results of the comparison of the theory with experimental data are shown in Tables 3 and 4 and in Figures 50 and 51.

One complicating circumstance must, however, be discussed in this case. The functions (11) and (12) are sym-

TABLE 3

EXPERIMENTAL DATA

S (Grams)	P_w	P_e	P_c
84.....	0.012	0.027	0.961
	.020		.980
88.....	.021	.082	.897
	(.053)		.947
92.....	.096	(.181)	.723
	.185		.815
96.....	.275	.266	.459
	.420		.580
100.....	.502	.267	.231
	(.683)		.317
104.....	.842	.103	.055
	.920		.080
108.....	.915	0.065	.020
	0.963		0.037

TABLE 4

THEORETICAL VALUES

S (Grams)	P_w	P_e	P_c
84.....	0.0039	0.0207	0.9754
	.0103		.9897
88.....	.0248	.0774	.8978
	(.0530)		.9470
92.....	.1027	(.1811)	.7162
	.1791		.8209
96.....	.2845	.2646	.4509
	.4092		.5908
100.....	.5512	.2502	.1986
	(.6830)		.3170
104.....	.7957	.1401	.0642
	.8797		.1203
108.....	.9320	0.0545	.0135
	0.9663		0.0337

metric with respect to $x = 0$. This means that a positive fluctuation is as likely to occur as a negative one. This occurs physically when there is no bias of any sort in favor of one of the stimuli. Frequently, however, a bias is introduced by the experimental setup. In the experiments of J. P. Guilford,[4] represented in Table 1, the two weights to be compared were presented in different orders several times. In the experiments of F. M. Urban,[5] represented in Table 3, the standard

weight was always presented first. This results, in the case of Urban, in a bias of the subject in favor of the second weight, which appears to be larger. Such a bias has the same effect as

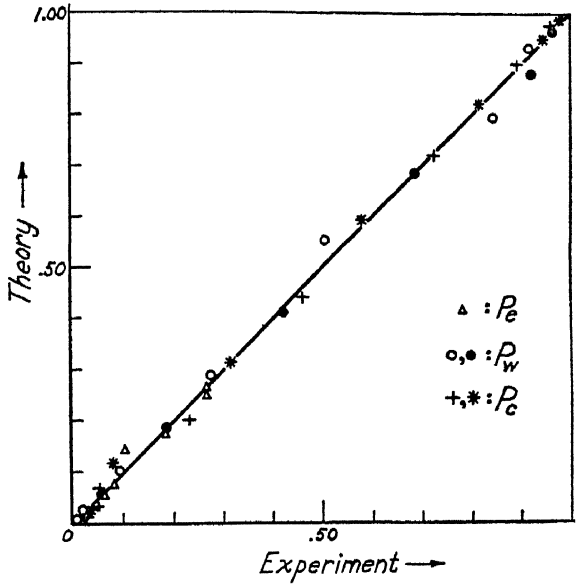


FIG. 50.—Same as Fig. 48 but for a different set of data, as given in Tables 3 and 4. Computed by H. D. Landahl.

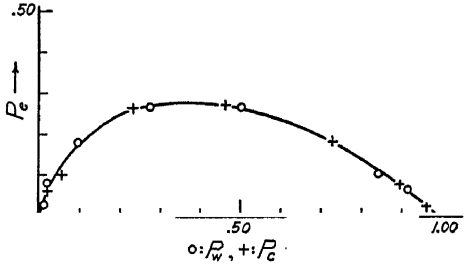


FIG. 51.—Same as Fig. 49 but for a different set of data. Computed by H. D. Landahl.

shifting the distribution function to the right or to the left along the axis of the abscissae. This gives, instead of (11),

$$p(x) = \frac{1}{\sqrt{2\pi\sigma}} e^{-(x-x_0)^2/2\sigma^2}, \qquad (16)$$

where x_0 is another parameter. A corresponding modification is introduced into function (12). Since now the distribution function is a two-parametric one, we have altogether three parameters: h, σ, and x_0. Therefore, three values from the experimental Table 3 are now used in order to calculate the remaining ones.

Another difference to be noted between the data of J. P. Guilford and those of F. M. Urban is that, in the former, all observations were made on one subject, while, in the latter, seven subjects were used, so that Table 3 is obtained by averaging for all subjects.

Analysis of visual and auditory data have also been made by H. D. Landahl.[2] He has also discussed more complex cases of more than two stimuli, as well as a relation between the probability of a correct response, the difference between the stimuli, and the time allowed for comparison of the stimuli.[1] This leads us into the interesting problem of the relation between the performance, difficulty of a task, and time allowed. For details we must refer the reader to the original papers.

REFERENCES

1. H. D. Landahl, *Psychometrika*, 3, 107, 1938.

2. H. D. Landahl, *Bull. Math. Biophysics*, 1, 159, 1939.

3. R. A. Fisher, *Statistical Methods for Research Workers* (Edinburgh and London: Oliver & Boyd, 1938).

4. J. P. Guilford, *Psychometric Methods*, pp. 187 and 195; Tables 28 and 30 (New York and London: McGraw-Hill Book Co., 1936).

5. F. M. Urban, *The Application of Statistical Methods to Problems of Psychophysics* (Philadelphia: Psychological Clinic Press, 1908).

CHAPTER XII

PERCEPTION OF VISUAL PATTERNS

As we have mentioned in chapter viii, when a constant stimulus is suddenly applied to a fiber characterized by

$$B < A\,; \quad b < a\,; \quad \frac{A}{a} < \frac{B}{b}\,, \tag{1}$$

then at first, for a short time, ϵ exceeds j; finally, however, j again exceeds ϵ and remains larger than ϵ as long as the stimulus lasts. This is due to the fact that, because of $A/a < B/b$, the asymptotic value AE/a of ϵ is less than the asymptotic value BE/b of j. Yet, because of $B < A$ and $b < a$, the quantity ϵ increases more rapidly at first, and also approaches its asymptotic value more rapidly than the quantity j. If, while the first stimulus S_0 lasts and j is therefore still larger than ϵ, we again apply, *suddenly*, an additional stimulus S_1, so that the total stimulus now becomes $S_0 + S_1$, then, as can be readily seen from Figure 52, again for a short time there will be excitation, followed again by a lasting inhibition (*MB*, chap. xxvi). It is, however, necessary that the additional stimulus be applied sufficiently suddenly (*MB*, chap. xxvi) and that it exceeds a certain threshold.

When such a fiber forms a synapse with another fiber, then the latter becomes excited for a short time only when the intensity of excitation of the first fiber *suddenly increases*. If the first fiber is excited *continuously* with a constant intensity E, the second fiber is not excited. As we have seen in *MB*, the intensity of excitation of the second fiber increases with increasing ΔE where ΔE is the amount of sudden increment

of E. As can be readily seen from the discussion in MB, chapter xxvi, the liminal ΔE which still produces an excitation of the second fiber is a linearly increasing function of E. Since E is, in general, a function $E = f(S)$ of the intensity of the peripheral stimulus, which involves the first fiber, therefore this liminal ΔE will, in general, be a more or less complicated function of S, depending on the shape of $f(S)$.

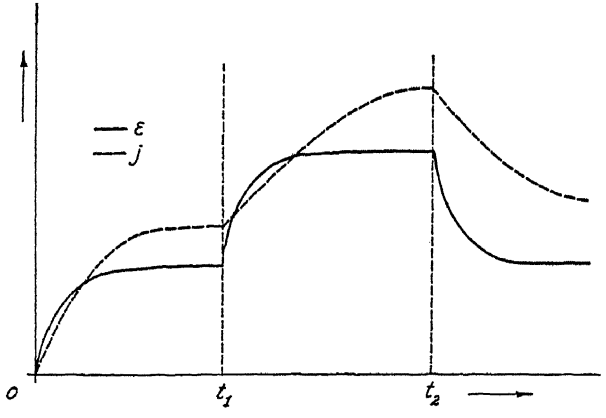


FIG. 52

If the second fiber leads to a region C_1, discussed in chapter x, then to each intensity of excitation of the second fiber there will correspond a definite group of excited neurons ne_3. And, since a definite intensity E^* of excitation of the second fiber corresponds to a definite relative sudden increment $\Delta E/E$ of the intensity of excitation of the first fiber, while this relative increment itself corresponds to a definite relative increment $\Delta S/S$ of the peripheral stimulus, therefore, to every value of $\Delta S/S$ there will correspond a definite individual group of neurons ne_3. As has been shown in MB, chapter xxvi, a similar mechanism is obtained for a *sudden* decrease of E or of S—that is, for a negative $\Delta S/S$.

We thus have a situation in which a *sudden change* of intensity of a constant external stimulus results in a short excitation of a definite group of central neurons, different groups corresponding to different amounts of change.

Combining this result with that of chapter x, we may say that to each absolute value of intensity of a peripheral stimulus there is a corresponding excitation of a definite center consisting of a definite group of individual neurons ne_3; and to each value of relative change of the intensity of peripheral excitation there is also a corresponding excitation of a similar definite center. The finite, though very large number of, neurons in the brain causes an overlapping of some of these centers and thus causes the perception of continuity of the possible values of intensity of peripheral excitation.

We shall now introduce some neurophysiological assumptions which may appear rather questionable. We are introducing them here, not because we consider them to be any more likely than other possible assumptions, but in a tentative way, as a sort of working hypothesis, in order to show how, by means of such or similar assumptions, we may develop a physicomathematical theory of some phenomena of visual perception, the discussion of which hitherto frequently remained on a purely qualitative level. Later on we shall discuss some possible modifications of these hypotheses—modifications which, however, may leave the fundamental formal results unaltered.

When we look at a segment of a straight line, we successively fix our attention on different points of this line. This results in movement of our eyes along that segment, back and forth. Experiments on eye movements[1] show that there is no simple relation between the shape of a figure contemplated and the path of the eye movement. Actual following of a rectangle gives a rather irregular path for the eye movement.[2] However, inasmuch as the eye muscles are innervated by

several centers in a rather complicated fashion, the following hypothesis will not necessarily be at variance with observations. Suppose that there is a group of brain centers which innervate the eye muscles in such a way as to make the eye follow *exactly* a contour which is projected on the retina. Each eye muscle is, however, innervated by a number of other centers, which produce different movements, superimposed on the above-mentioned movements. Thus, the actual move-

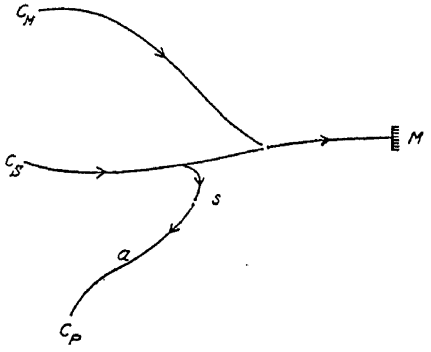


FIG. 53

ment of the eye may have no apparent relation to the shape of the contour contemplated; yet it contains *as a part* those movements which trace an exact replica of the geometrical pattern looked at.

In Figure 53, C_S represents schematically the center which would produce such exact movements of the eye, through the muscle M, while C_M stands for all other centers which may impart to M movements that have nothing to do with the shape of the contour. The character of the movement which would be produced by M under the influence of C_S alone depends on the character of excitation of C_S (intensity, duration). If through a synapse s a fiber a is excited which leads to a sensory center C_P (Fig. 53), then to each definite type of

excitation of C_S, or, in other words, to each definite type of movement produced by M under the influence of C_S alone, there corresponds a definite intensity and duration of excitation of C_P. If M were to be innervated by C_S only, then the same type of excitation in C_P could be obtained through proprioceptive fibers, which would lead from M to C_P and which would be excited by the contraction of M. In such a case the proprioceptive impulses would have a definite relation to the shape of the contemplated pattern. Actually, because of the other centers C_M, the proprioceptive impulses coming from M have no relation to the shape of the pattern. But the impulses coming through s over a (Fig. 53) do have such a relation. In other words, according to the scheme represented by Figure 53, things will occur at C_P *as if* the movements produced by M would exactly correspond to the shape of the contour contemplated and *as if* these movements would send off proprioceptive impulses directly to C_P.

The considerations above therefore enable us to speak, for simplicity, of the eye as actually following the contemplated contour, and of proprioceptive impulses *corresponding to such movements* being sent to some sensory centers. Although actually things are much more complex, by means of this scheme we may interpret our schematized statements in terms of a more exact neurophysiological picture. Thus, when in the following we say: "The contraction of the eye muscle sets up proprioceptive impulses which behave so and so," we do not mean it literally, but actually imagine a scheme like the one represented by Figure 53.

Neurological mechanisms that would produce such a movement can be readily suggested. For the time being, let us make, for simplicity and as a special, purely theoretical case, the assumption that this movement of the eyes goes on with constant velocity v. Let a line segment AB be placed in a vertical plane, and let its angle with the vertical be θ (Fig.

54). Neglecting the possible role of the oblique eye muscles, the movement of the eyes in the direction AB involves a contraction of the rectus superior of both eyes—the right rectus lateralis externus and the left rectus lateralis internus—and a relaxation of both recti inferiors—the right lateralis internus and left lateralis externus.[1]

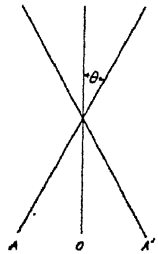


FIG. 54

Again assuming for simplicity that the coordinated movements of the left lateralis internus and the right lateralis externus are controlled from one common center, and that the same holds for the right lateralis internus and the left lateralis externus, we may greatly simplify our considerations by considering only the movements of *one eye*—for instance, the right one. Then the movement in the direction AB (Fig. 54) involves a contraction of rectus superior and of lateralis externus and a relaxation of rectus inferior and lateralis internus which may be written in an abbreviated form, thus:

$$AB{:}rs+ \qquad le+ \qquad ri- \qquad li- \tag{2}$$

The movement in the direction BA involves

$$BA{:}rs- \qquad le- \qquad ri+ \qquad li+ \tag{3}$$

The changes of muscular contraction of the four muscles produce proprioceptive impulses of different intensities; and if the proprioceptive fibers lead to a region C_1, discussed in chapter x, then any given intensity of contraction will produce excitation in a definite individual group of neurons ne_3. Again, for simplicity, we shall consider that only positive contractions produce proprioceptive impulses, a relaxation being characterized by a lack of such impulses. Let the intensity of these impulses be proportional to the *velocity* of the con-

traction of the muscles. Then, if v is the velocity of the eye movement, the velocities of contraction of rs and le in (2) are, respectively,

$$v \cos\theta \quad \text{and} \quad v \sin\theta\,;$$

and the corresponding intensities of the proprioceptive impulses are

$$E_{rs} = av \cos\theta\,; \qquad E_{le} = av \sin\theta\,, \tag{4}$$

where a is a coefficient of proportionality.

The excitation E_{rs} excites a definite individual group of neurons ne_3 in a center which we shall call V (for vertical); while E_{le} excites a group of neurons in a center H (horizontal).

The movement of the eye in the direction BA produces the following excitation:

$$E_{ri} = av \cos\theta\,; \qquad E_{li} = av \sin\theta\,, \tag{5}$$

and causes the excitation of two other groups of neurons ne_3, one in V and one in H. If the eye moves up and down indefinitely, four groups of neurons are excited. If we now consider another segment of a straight line, which is characterized by a different $\theta = \theta'$, again four groups of neurons ne_3 will be excited; but these four groups will be different from the former four, because E_{rs}, E_{le}, E_{li}, and E_{ri} are different. Some of the new groups may partly overlap the old ones only when $\theta - \theta' = \Delta\theta$ is very small. Thus, every position of a straight-line segment in a vertical plane corresponds to the excitation of four distinct individual centers, two in H and two in V.

Considering, now, the segment $A'B'$ (Fig. 54), symmetrical to AB with respect to the vertical OO', we shall find, by a similar argument, that the scanning of that segment with the eye in the direction $A'B'$ gives

$$E'_{rs} = av \cos\theta = E_{rs}\,; \qquad E'_{li} = av \sin\theta = E_{li}\,, \tag{6}$$

and in the direction $B'A'$ gives

$$E'_{ri} = av \cos\theta = E_{ri}\,; \qquad E'_{le} = av \sin\theta = E_{le}\,. \tag{7}$$

That is, AB and $A'B'$, scanned in both directions, *will excite identical groups of neurons* ne_3 *in* H *and* V. This difficulty can be avoided in several ways, only one of which we shall consider here.

Let, as suggested by Gale Young, the neurons of higher order, corresponding to *li*, *le*, and *ri*, have such high thresholds that they remain unexcited no matter how strong the excitation of the peripheral fibers. However, let those higher-order neurons also be excited by collaterals from *rs*, so that while, for instance, excitation of *li* alone does not result in any central excitation, yet simultaneous excitation of *rs* and *li* results in such a central excitation.

In that case a contraction of the muscle *rs*, that is, a movement of the eye upward, if accompanied by a lateral movement, produces an excitation both in *V* and *H*. But a downward movement of the eye, produced by a contraction of *ri* and a relaxation of *rs*, does not result in any excitation in either *V* or *H*. Of course, other brain centers may be excited by such movement by means of other fiber connections. But at present we are interested only in *V* and *H*. We now have, for the movement along AB (Fig. 54),

$$AB{:}rs+\,; \qquad le+\,; \qquad ri-\,; \qquad li-\,; \tag{8}$$

and for BA:

$$BA{:}rs-\,; \qquad le-\,; \qquad ri-\,; \qquad li-\,. \tag{9}$$

For $A'B'$ we have

$$A'B'{:}rs+\,; \qquad li+\,; \qquad ri-\,; \qquad le- \tag{10}$$

and

$$B'A' : rs - ; \quad li - ; \quad ri - ; \quad le - . \tag{11}$$

From (8) and (9) it follows that scanning of AB in both directions results in

$$E_{rs} = av \cos\theta ; \quad E_{le} = av \sin\theta , \tag{12}$$

while scanning $A'B'$ in both directions results in

$$E'_{rs} = av \cos\theta ; \quad E'_{li} = av \sin\theta . \tag{13}$$

Both AB and $A'B'$ involve rs in the same amount and therefore excite the same group of neurons in V. But, while AB involves le, $A'B'$ involves li. Hence, the two groups on neurons in the H-center, which correspond to the scanning of the two lines in both directions, will be different. Every segment of the straight line excites one group of neurons in H and one in V, the groups differing for different θ's. Lines symmetrical with respect to either a vertical or a horizontal line have a common group in V. Parallel segments produce identical excitation.

In order to obtain excitation in H center when a horizontal line is scanned, not involving any upward movement of the eye, we must introduce an additional assumption, namely, that only an actual downward movement of the eye results in a relaxation of rs, while holding the eye on the same level still requires a tonic contraction of rs.

The proprioceptive excitations, owing to the contemplation of any segment of a straight line, may thus, through V and H, be conditioned to a response R. A different segment will, in general, not produce R, because it involves different groups of neurons in V and H. A segment symmetrical to the original one with respect to a vertical line may, however, pro-

duce a weaker R through the group of neurons in V, which it has in common with the original segment.

Now consider two segments, AB and BC (Fig. 55), forming an angle at B. Following this figure with the eye results in the excitation of two groups of neurons in V and of two in H. If, however, the proprioceptive fibers, leading from rs, li, and le to V and H, send off collaterals, which are characterized by relations (1) and each of which leads to a center A^*, A^* will become excited every time the eye passes the angle B, because of a sudden change in the intensity of excitation E_{rs} and E_{le} (or of E_{rs} and E_{li} for a differently oriented angle).

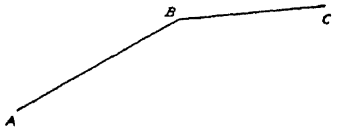


FIG. 55

To each value of the angle B there is a corresponding different intensity of excitation in A^*; and if A^* is connected to a center A, of the kind discussed in chapter x, then to each value of the angle B there corresponds a definite group of neurons ne_3 in A. We thus have a center for the perception of angles.

Making different hypotheses concerning the physical characteristics and constants of all centers involved, we may derive various quantitative relations which may suggest experimental studies. As an illustration, an expression has been derived for the relation between the minimum perceptible angle formed by the two straight lines AB and BO and the angle which one of them—say AB—forms with the vertical.[3]

Finally, we may consider not only that the speeds of the contraction of rs, li, and le control the excitation of some centers but that special fibers are provided in which the inten-

sity of excitation is proportional to the amplitude of the muscular movements. In this fashion, by considerations similar to those above, we shall arrive at a picture in which a particular group of neurons ne_3 is excited in a center L, the group being characteristic for the length of the segment of straight line considered.

If a polygonal contour, consisting of n sides and n angles, is presented to the subject, then it follows from the foregoing that, as this contour is followed by the eye, in general, n groups of neurons will be excited in V, n groups in H, n groups in A, and n groups in L. The intensities of excitation of the different groups, even belonging to the same center, such as H or V, will generally be different. Two groups of neurons ne_3, corresponding to two different distinct intensities E, will not only be distinct but will, in general, contain different numbers of neurons. If the polygon possesses some symmetry properties, which result, for instance, in some of the angles being equal, this reduces the number of distinct groups in A. If there are m equal sides, then they all excite the same group of neurons ne_3 in L, etc. As the contour is followed by the eye, the excitation of each group comes and goes, discontinuously, occurring only when the attention is concentrated on the particular element (segment, angle, or length). If, out of n elements, m are identical, the group ne_3 corresponding to the m identical elements k, will be excited, on the average per unit time, m times more often than those groups ne_3 which correspond to elements that have no other identical ones. Hence, *the average intensity of excitation* of the group corresponding to m identical elements k will be m times the excitation due to a single element k.

We may now consider a scheme, discussed in *MB*, chapter xxii, and in chapter x of this present work, in which an excitatory fiber leads from neurons of each center H, V, A, and L to neurons of higher corresponding centers, H', V', A',

and L'. This excitatory fiber branches off into inhibitory fibers, leading to the neurons of all other centers. We may ask for the total excitation corresponding to a given polygonal contour. If this total excitation is transmitted to a center whose excitation results in a sensation of pleasure (MB, chap. xxx), then its intensity may be considered as a measure of the pleasantness or of the aesthetic value of a given contour. The problem is perfectly definite and can be treated

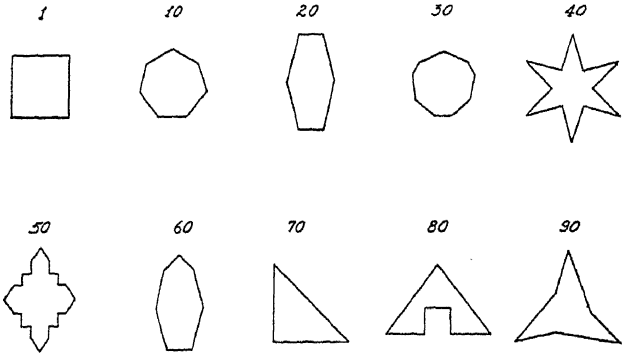


FIG. 56.—These ten polygons, selected from one hundred given in G. Birkhoff's book,[4] were used in the experiments of R. C. Davis[5] to determine the relative aesthetic values of different geometric patterns.

provided the intensities of excitation of all the groups ne_3 in H, V, A, and L are given—in other words, provided the distribution function $N(h)$, discussed in chapter x, page 149, for different centers, as well as the function given by equation (1) of chapter x, is specified. For different choices of these functions the problem will be of different degree of mathematical difficulty. A particularly simple case is obtained when we consider a very special case where the intervals (h_1, h_2) (cf. chap. x, Fig. 39) always contain the same number of neurons and where the intensity of excitation of any such group, falling in any interval (h_1, h_2), is approximately con-

stant. Such a case is rather unlikely to occur exactly, but we shall consider it here as a possible approximation. In that case two different elements—for instance, angles—are characterized by two distinct groups ne_3, having, however, the same total intensity of excitation E_0 when the element is perceived. To m identical elements there corresponds one group with an *average* intensity mE_0, since the group is stimulated m times more frequently.

We shall apply the foregoing considerations to a group of polygons for which G. Birkhoff[4] has calculated quantitative aesthetic values on the basis of some general considerations of complexity and order and for which those values have been measured by R. C. Davis by the rank-order method[5] (Fig. 56). Let the intensity E_0 of excitation of a group ne_3 be taken as 0.1. The intensity of inhibition, which is approximately proportional to E_0 (chaps. ix and xi) is taken to be 0.01 per unit of E_0. Then for No. 50 of Figure 56, represented separately on Figure 57, we have the following:

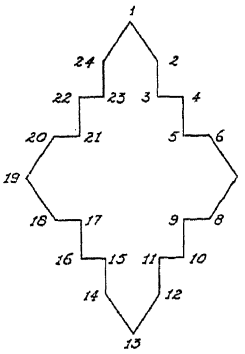


FIG. 57

There are, altogether, eight segments of horizontal lines which do not give any excitation at V: 3–4, 5–6, 8–9, 10–11, 15–16, 17–18, 20–21, and 22–23; they all cause excitation of one group ne_3 in H, with an intensity of 0.8. The sides 1–2, 6–7, 13–14, and 18–19 are all parallel and therefore identical in their H and V groups. Hence, in the H center they give one group with $E_0 = 0.4$. Similarly, 1–24, 19–20, 12–13, and 7–8 give, in the H center, one group with $E_0 = 0.4$. In the V center the sides 2–3, 4–5, 9–10, 11–12, 14–15, 16–17, 21–22, and 23–24 give one group with $E_0 = 0.8$; while 1–2, 6–7, 7–8, 12–13, 13–14, 18–19, 19–20, and 24–1 give another group

in V with $E_0 = 0.8$. In the A center we have one group with $E_0 = 1.2$ corresponding to the right angles 3, 4, 5, 9, 10, 11, 15, 16, 17, 21, 22, and 23; a group with $E_0 = 0.2$ for 1 and 13; another group with $E_0 = 0.2$ for 7 and 19; one group with $E_0 = 0.4$ for 2, 12, 14, and 24; and one group with $E_0 = 0.4$ for 6, 8, 18, and 20. In the L center we have one group with $E_0 = 0.8$ for the eight equal sides 2–3, 4–5, 9–10, 11–12, 14–15, 16–17, 21–22, and 23–24; another group with $E_0 = 0.8$ for the eight equal horizontal sides; and a third group with $E_0 = 0.8$ for the eight equal inclined sides. Altogether we have the following scheme:

H	1 group	of $E_0 = 0.8$
	2 groups	of $E_0 = 0.4$
V	2 groups	of $E_0 = 0.8$
A	1 group	of $E_0 = 1.2$
	2 groups	of $E_0 = 0.2$
	2 groups	of $E_0 = 0.4$
L	3 groups	of $E_0 = 0.8$

To each group in the centers H, V, A, and L there is a corresponding group in the centers H', V', A', and L'. The first group of H' receives inhibitory fibers from all other groups, that is, from twelve groups altogether. The corresponding intensities of inhibition are 0.004, 0.004, 0.008, 0.008, 0.012, 0.002, 0.002, 0.004, 0.004, 0.008, 0.008, and 0.008—altogether 0.072. Hence, the net excitation of the first group in H' is $0.8 - 0.072 = 0.728$. Each of the two other groups in H' again receives inhibitory fibers from twelve other groups, and the net excitation of the group is thus found, in a similar way, to be 0.324. Calculating, thus, the net excitation for each group and adding them all up, we obtain, for the total value of excitation, 7.04. In a similar way the total intensity of excitation of the groups were computed for the other ten polygons of Figure 56, and the results are plotted on Figure 58

(full light line). The broken and alternate lines of Figure 58 represent the experimental rank-order values of R. C. Davis for two different groups of students. The dotted line represents values calculated by Birkhoff. Inasmuch as our values

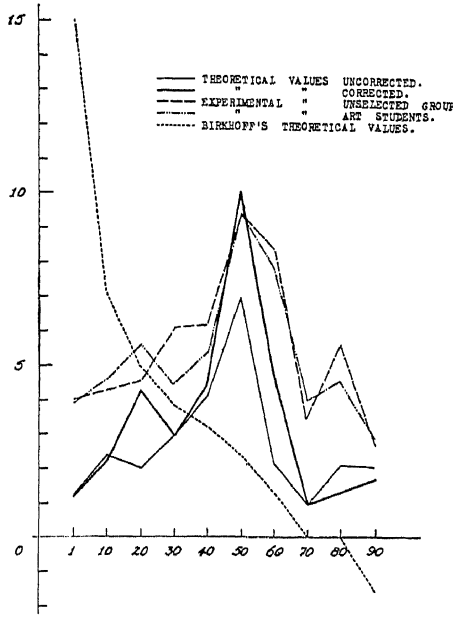


FIG. 58.—The two full lines represent the theoretical values of the intensity of aesthetic sensation in arbitrary units for the ten polygons represented in Fig. 56. The broken and the alternate lines represent the experimental data by R. C. Davis.[5] The dotted line represents theoretical values calculated by G. Birkhoff.[4] See text, pp. 189 ff.

give the actual intensities of central excitation, measured in arbitrary units, while Davis' values give only the rank order, in which a certain arbitrariness enters in the assignments of weights, we cannot expect a close numerical agreement between our values and those of Davis. However, the general trend agrees, with the exception of No. 20.

We notice that for No. 20 and No. 60 our theory gives values that are too low. The fact that No. 60 has actually a much higher value than No. 10, although both are septagons with approximately similar symmetry properties, suggests that the V-groups, which correspond to smaller θ, are more strongly excited than those corresponding to larger θ's. The assumption of unequal excitation of the different groups would make our calculations rather cumbersome. Leaving this more rigorous study for the future, we shall approximately take into account a possible preference for vertical directions by multiplying the values of total intensity of excitation, as calculated, by the ratio of the heights to widths of the corresponding polygons. This is obviously a crude method, holding only for ratios which are not too large. The results are shown in Figure 58, heavy line. The general agreement is somewhat improved, at the expense of No. 80, which is too much lowered. In view of the very crude approximations made, these results should command our attention. Further refinement and generalization of the theory, along lines indicated above, is likely to bring further improvements.

If we do not make the rather crude assumption that all the centers are excited with equal intensity, then, obviously, the final results will depend on the assumption about the intensity of excitation of the different centers. These intensities of excitation may vary from individual to individual, thus accounting for the differences in tastes (*MB*, chap. xxix). On the average, these intensities may be approximately equal, which would account for the fair agreement of the present theory with observations made on large groups of people.

The influence of various symmetries, that is, the appearance of several identical elements, may be easily analyzed by considerations similar to those given in chapter xxii of *MB*. If we have, for instance, n equally externally excited but

different elements, then each of the corresponding groups ne_3 receives an inhibition from $n - 1$ other groups. If m of these elements are identical, then we have only $n - m + 1$ groups. The total applied excitation remains the same. In the first case it was nE; now it is $(n - m)E + mE$. But the group corresponding to the m identical elements receives inhibitory fibers from $n - m$ other groups. Thus the total inhibition is decreased, and the net excitation increased. If we take, as a measure of complexity, the total number of elements in the pattern, then various types of symmetry have the effect of reducing the "effective" complexity; but this reduction occurs by a process of subtraction rather than by a process of division, as in Birkhoff's theory. It must also be remarked that, even when all elements are different, the total excitation in our case does not monotonically decrease with the complexity n. Equation (1) of chapter xxix of *MB* gives us the intensity of excitatory process at one of the n synapses, corresponding to n different elements with equal intensity of excitation. The total intensity of excitation is obtained by multiplying that equation by n, which gives

$$E_{\text{tot}} = (K_1S - K_4)n - (K_2S - K_3)n(n - 1) .$$

This expression is zero for $n = 0$ and has a maximum for

$$n^* = \frac{(K_1 + K_2)S - (K_3 + K_4)}{2(K_2S - K_3)} ,$$

after which it decreases to zero and becomes negative. We should therefore expect that in a rather simple pattern, consisting of a very few elements, too much symmetry will be rather a disadvantage from the aesthetic point of view, for in that case the total excitation increases with increasing complexity. But for very complex patterns, where the complexity is greater than the optimal n^*, some symmetry is pleasant, as

reducing the complexity and bringing it closer to the optimal. In the general case of unequally excited elements, these relations become much more complicated.

If a subject or animal is conditioned to a certain polygonal figure and then he is presented with another figure, which has some elements in common with the first one, the intensity of the response to the second figure can be calculated by the method used in chapter xxix of *MB*. Thus, a quantitative theory of discrimination of patterns may be developed. Two patterns will be the more similar the more they have common groups ne_3 in various centers. If all the neurons in A send off fibers to a center A_0, so that A_0 is always excited, whenever any angle, no matter of what size is presented, then A_0 is a center that responds to "angularity" in general. An animal trained to choose a square, when a square and a circle are presented, will choose the more angular figure when two other, but different, figures are presented. This provides a basis for the understanding of some interesting results of H. Klüver[6] on the equivalence or nonequivalence of certain pairs of geometrical figures in the behavior of monkeys.

The mechanism discussed here suggests a definite procedure for the generalization of the theory, including the aesthetic measure to curvilinear figures. If we again consider a fiber of the kind discussed on page 177, but if we consider, instead of a sudden change of E, a gradually increasing E, then we shall find, by using the fundamental equations developed in *MB*, chapter xxii, that to any given variation of E with respect to time there will correspond a definite intensity of excitation of a definite duration. When a curve is followed by the eye with a definite velocity, then E_{rs}, E_{li}, and E_{le} are not constants but are known functions of time, determined by the shape of the curve. Therefore, we may calculate the corresponding intensity of the excitation in a center C_c, to which appropriate fibers are leading. We may

then calculate the minimum perceptible curvature and other interesting problems. We may also ask what particular shape of curve gives the maximum excitation at C_c (cf. *MB*, chap. xxii).

Experiments by Edmund Jacobson[7] indicate that thinking of a vertical object, while not causing actual muscular contractions, sets up very weak action currents in the eye muscles, such as would correspond to a vertical movement of the eye. Thinking of horizontal objects gives currents corresponding to horizontal movements. This suggests also that proprioceptive impulses of the kind discussed on page 180 may be set forth without the muscles performing the actual movements. It has been observed that complete paralysis of the eye muscles does not affect recognition of shapes. This may be due to the taking-up of the functions of the eye muscles by other muscles of the body, which may imperceptibly move as a whole; or it may involve a mechanism of Gestalt recognition, discussed previously (*MB*, chap. xxvii).

Finally, it may be shown[8] that a mechanism based on our fundamental equations developed in *MB* can be conceived in which, when a chain of neurons is excited along a contour line in a brain center, a wave of excitation is set up, moving along that chain with constant velocity. The vertical and horizontal component of the velocity of that wave may then formally play the role of the vertical and horizontal movements of the eye. If such a mechanism is at work, then the central wave may continue to travel even for some time after the visual object is removed from the field of vision.

The main thing is that we have here a picture which provides for the analysis of complex patterns and also gives a basis for a quantitative, exact theory. Most likely this mechanism is operating simultaneously with the one discussed before (*MB*, chap. xxvii). The latter provides for recognition of gross features and perception of "wholeness"; the former, for analysis of details.

REFERENCES

1. A. Tachermak, "Augenbewegungen," *Handbuch d. norm. und path. Physiologie*, 12, Part II, 1001 (Berlin: J. Springer, 1931).

2. *Ibid.*, p. 1059.

3. N. Rashevsky, *Psychometrika*, 3, 253, 1938.

4. G. Birkhoff, *Aesthetic Measure* (Cambridge, Mass.: Harvard University Press, 1933).

5. R. C. Davis, "An Evaluation and Test of Birkhoff's Aesthetic Measure Formula," *Jour. Gen. Psychol.*, 15, 231, 1936.

6. H. Klüver, *Behaviour Mechanisms in Monkeys* (Chicago: University of Chicago Press, 1933).

7. E. Jacobson, "Electrical Measurements of Neuro-muscular States during Mental Activity. III. Visual Imagination and Recollection," *Amer. Jour. Physiol.*, 95, 694–702, 1930.

8. H. D. Landahl, *Bull. Math. Biophysics*, 1, 95, 1939.

INDEX